Center for Advanced Computation

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, ILLINOIS 61801

CAC Document No. 52

SOME EXTREMAL PROBLEMS IN GEOMETRY AND ANALYSIS

By

George B. Purdy

July 1, 1972

CAC Document No. 52

SOME EXTREMAL PROBLEMS IN GEOMETRY AND ANALYSIS

By

George B. Purdy

July 1, 1972

Center for Advanced Computation
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

Abstract

This report treats three optimization or extremal problems in geometry and analysis. In the first chapter we consider the problem of finding positive minima of certain Dirichlet L-functions with positive real arguments. Optimal triple lattice packing of spheres of equal radius is the subject of the second chapter. Packing and covering problems with spheres are important in information theory for the design of communication networks. In the last chapter we treat an extremal problem about triangles of equal area.

TABLE OF CONTENTS

I. THE REAL ZEROS OF THE EPSTEIN ZETA FUNCTION

§1. Introduction If $d > 4$ and $-d$ is a fundamental discriminant -- that is, of $-d \equiv 1,5,8,9,12$ or 13 (modulo 16) and $-d$ is not divisible by the square of any odd prime, then we define $L_{-d}(s) = \sum_{n=1}^{\infty} (-d|n)\, n^{-s}$ (Re $s > 0$), where $(-d|n)$ is the Kronecker symbol. M. E. Low proved in Theorem 2 of [1] that a certain hypothesis, which we will call Low's Hypothesis, implied $L_{-d}(s) > 0$ for $s > 0$. Low then put an IBM 7094 to work to check this hypothesis for small values of d and concluded that if $-d$ is a fundamental discriminant and $d < 593{,}000$, then $L_{-d}(s) > 0$ for $s > 0$ with the possible exception of $L_{-115147}$.

G. J. Rieger, who reviewed Low's paper in the Mathematical Reviews, asserted that Low's results were doubtful because he did no error analysis of his computation. Dr. Paul Bateman suggested to me that I do a computer verification of Low's Hypothesis using higher precision arithmetic than Low did, extend the range, and at the same time do an analysis of the error to ensure the validity of the results.

We tested Low's Hypothesis for $d \leqq 800{,}000$ on an IBM 360/75 using double precision arithmetic and discovered one counterexample overlooked by Low -- namely $d = 357819$. Low's Hypothesis is true for $1 \leqq d \leqq 800{,}000$ with three exceptions.

We do not wish to quote all the details of [1] here. We give the definition of Low's Hypothesis and refer the reader to [1] for the proof that it implies $L_{-d}(s) > 0$ for $s > 0$.

Low's Hypothesis (i) $\sum_Q \alpha(a,d)/\sqrt{a} \geqq 0$

(ii) $\sum_Q \{\alpha(a,d) - H_0(a,b,c)\}/\sqrt{a} \geqq 0$

(iii) $\sum_Q \{b_{2n+1} + \alpha^{2n+1}(a,d)/(2n+1)!\}/\sqrt{a} \geqq 0$ $(n \geqq 1)$.

Here

(1) $\alpha(a,d) = \log a + \log(8\pi e^{-\gamma}) - (1/2)\log d,$

$b_n = (2^n-1)\zeta(n)/n + 2^n\beta_n,$

and β_n is given by

$\log\{(s-1)\zeta(s)\} = \sum_{n=1}^{\infty}(-1)^n\beta_n(s-1)^n \qquad (|s| < 3),$

and Q is the set of triples (a,b,c) where a,b,c are the coefficients of a reduced quadratic form of discriminant -d, and $H_o(a,b,c)$ is an upper bound on the error function, H(s; a,b,c) of Bateman's and Grosswald's approximation to the Epstein zeta function -- see [3]. More exactly,

$Q = \{(a,b,c)\colon -d = b^2-4ac,\ -a < b \leqq a < c \text{ or } 0 \leqq b \leqq a = c;\ a,b,c \text{ integers}\},$

and

(2) $H_o(a,b,c) = 2k^{-1/2}e^{-2\pi k}\cos(\pi b/a) + (0.08)k^{-1/2}e^{-2\pi k}|\cos(\pi b/a)|,$

where $k = \sqrt{d}/(2a)$.

It is easily seen -- see [1] -- that $H_o(a,b,c) \leqq 0.005$. Hence, the condition,

(3) $\sum_Q\{\alpha(a,d) - 0.005\}/\sqrt{a} \geqq 0,$

implies both (i) and (ii) of Low's Hypothesis. As written, condition (iii) is an infinite sequence of inequalities for n = 1,2,3 . . . However, the condition is trivially satisfied for large n. To see this, suppose that $d \leqq d_o$. Then $\alpha(a,d)$ is bounded and $b_{2n+1} > 6.34$ for $n \geqq 3$ -- for this see [1]. Hence $b_{2n+1} + \alpha^{2n+1}(a,d)/(2n+1)!$ will be positive for n sufficiently large, depending on d_o. Summing, we get (iii) for $n \geqq n_o(d_o)$. This leaves only finitely many sums to check. Since (i) and (ii) are finitely computable, we see that Low's Hypothesis is finitely computable for every d.

Low shows in [1] that (iii) is a consequence of (i) for $d < 593{,}000$. His proof is in three parts. First of all $\sum_Q\{b_3+\alpha^3(a,d)/6\}/\sqrt{a} > 2\sum_Q\alpha(a,d)/\sqrt{a}$ for $d < 593000$. Secondly, $\sum_Q\{b_5+\alpha^5(a,d)/120\}/\sqrt{a} > 2.5\sum_Q\alpha(a,d)/\sqrt{a}$ for $d < 1{,}970{,}000$, and finally $b_{2n+1}+\alpha^{2n+1}(a,d)/(2n+1)! > 0$ for $d < 1{,}320{,}000$, and $n \geq 3$.

It was therefore possible for us to prove (iii) for $d < 1{,}320{,}000$ by testing on the computer the condition

(4) $\quad \sum_Q\{b_3+\alpha^3(a,d)/6\}/\sqrt{a} > 0, \; 593000 \leq d < 1{,}320{,}000.$

The number $|Q|$ of triples (a,b,c) in Q is the class number of $-d$, denoted by $h(-d)$. We need the following estimate later:

(5) $\quad |Q| < 4000$ for $d \leq 80{,}000$.

This follows from the following result:

<u>Lemma 1</u> $h(-d) = (\sqrt{d}/\pi)\,(L_{-d}(1)) < (\sqrt{d}/\pi)\log d < 3900$ for $4 < d \leq 800{,}000$.

<u>Proof</u> The first equation is well known. Let $\chi(n) = (-d|n)$, the Kronecker symbol, and let

$$s(n) = \chi(1) + \chi(2) + \ldots + \chi(n).$$

Then

$$\begin{aligned} L_{-d}(1) &= \textstyle\sum_{n=1}^{\infty}\chi(n)/n \\ &= \textstyle\sum_{n=1}^{\infty}(s(n) - s(n-1))/n \\ &= \textstyle\sum_{n=1}^{\infty}s(n)(1/n - 1/(n+1)) \\ &= \textstyle\sum_{n=1}^{\infty}s(n)/(n(n+1)). \end{aligned}$$

Let $k = \max_n |s(n)|$. Then $k \leq (d-1)/2$, and

$$\begin{aligned} L_{-d}(1) &\leq \textstyle\sum_{n=1}^{k-1} 1/(n+1) + k\sum_{n=k}^{\infty}1/(n(n+1)) \\ &= \textstyle\sum_{n=1}^{k-1}1/(n+1) + k\sum_{n=k}^{\infty}(1/n - 1/(n+1)) \end{aligned}$$

$$= 1/2 + 1/3 + \dots + 1/k + 1$$

$$\leqq \int_{1/2}^{k+1/2} dx/x$$

$$= \log(2k+1) \leqq \log d.$$

Hence $h(-d) = (\sqrt{d}/\pi)L_{-d}(1) < \sqrt{d}(\log d)/\pi < (\sqrt{800{,}000}\,\log(800{,}000))/\pi < 3900$, for $4 < d < 800{,}000$.

§2. The Programs In testing (3) and (4) it is not necessary to evaluate the whole sum, because many of the terms are known to be positive. Let Q_1 consist of the triples (a,b,c) of Q such that $|b| < 128$ and let Q_2 consist of the others. Then

$$\sum_Q \{\alpha(a,d) - 0.005\}/\sqrt{a} = \sum_{Q_1} \{\alpha(a,d) - 0.005\}/\sqrt{a} + \sum_{Q_2} \{\alpha(a,d) - 0.005\}/\sqrt{a},$$

and the terms $\{\alpha(a,d) - 0.005\}/\sqrt{a}$ in the second sum are positive for $d < 3{,}000{,}000$. To see this we note that since $a \geqq |b| \geqq 128$, we have

$$\alpha(a,d) - 0.005 = \log a + \log(8\pi e^{-\gamma}) - 1/2 \log d - 0.005$$

$$> .035 > 0.$$

Two programs were used to verify (3). Program One tested the condition

$$(6) \qquad \sum_{Q_1} \{\alpha(a,d) - 0.005\}/\sqrt{a} \geqq 1/10$$

in double precision arithmetic for all $d \equiv 1,5,8,9,12,13$ (modulo 16) between 5 and 800,000 recording the failure (approximately 2%) onto magnetic tape along with the sum $\sum_{Q_1}$. The 1/10 allows margin for error due to round-off.

Program Two reads the tape and adds terms from $\sum_{Q_2}$ until the resulting sum exceeds 1/10; however if the resulting sum never exceeds 1/10, then d is again written on magnetic tape to be read by Program Three.

A third program, Program Three, verifies conditions (i) and (ii) directly for those few hundred d's failing Program Two. Only three numbers failed condition (i), and none failed (ii); the numbers were 115147, 357819 and 636184.

§3. The Error Analysis The calculations were performed using double precision floating point arithmetic on the IBM 360/75 at the University of Illinois. The numbers representable exactly in this way are of the form

$$r\ 16^{p-14},$$

where r and p are integers, $16^{13} \leqq |r| \leqq 16^{14}-1$, and $|p| < 64$.

Thus small integers and rational numbers with denominator 2^k for small k are representable exactly. Any other number x is approximated by $(1 + \varepsilon)x$ where $|\varepsilon| \leqq 16^{-13}$.

Following a notation introduced by Wilkinson, we shall use $f\ell(E)$ to denote the computed value for the mathematical expression E. When the order of the calculation is ambiguous, such as in $E = a+b+cd$, it is assumed that the calculation proceeds from left to right; thus $E = ((a+b) + cd)$.

For any reasonable machine,

$$f\ell(a) = (1 + \varepsilon_1)\ a$$

$$f\ell(a+b) = (1 + \varepsilon_2)\ f\ell(a) + (1 + \varepsilon_3)\ f\ell(b),$$

$$f\ell(a-b) = (1 + \varepsilon_4)\ f\ell(a) - (1 + \varepsilon_5)\ f\ell(b),$$

$$f\ell(a \times b) = (1 + \varepsilon_6)\ f\ell(a)\ f\ell(b),$$

and

$$f\ell(a/b) = (1 + \varepsilon_7)\ f\ell(a)/f\ell(b), \text{ where } |\varepsilon_i| \leqq \delta \quad \text{for } 1 \leqq i \leqq 7.$$

There could well be some discussion about the value of δ. If the machine 360/75 were designed by a numerical analyst, we could assume that $\delta \leqq 16^{-13} < 2.23 \times 10^{-16}$, but machines today usually are not. We shall see however that δ does not have to be anywhere near that small for our calculations to be valid - e.g. $\delta = 10^{-12}$ would be adequate.

The mathematical functions $\cos x$, e^x, $\sqrt{x}$, and $\log x$ were calculated using the double precision routines in the Fortran Library; we need the following:

$$fl(\cos x) = \cos x + \theta_1 \qquad (0 \leq x \leq \pi),$$

$$fl(e^x) = (1 + \theta_2)\, e^x + \theta_3 \qquad (-100 \leq x \leq 0),$$

$$fl(\sqrt{x}) = (1 + \theta_4)\, \sqrt{x} \qquad (1 \leq x \leq 1000),$$

and

$$fl(\log k) = (1 + \theta_5) \log k \qquad (k = 1, 2, \ldots, 10^6),$$ where

$|\theta_i| \leq M$, and M is not too large.

We do not know whether anyone has done an analytical error analysis of these routines. Perhaps Dr. Kuki of the University of Chicago (now deceased), who designed the routines did one. Kenneth E. Hillstrom of Argonne National Laboratory -- see [4] -- did a test on these routines, using random arguments over several intervals and comparing the results with higher precision calculations on another machine. His results suggest $|\theta_i| \leq 1.7 \times 10^{-13}$ for $1 \leq i \leq 4$, but he naturally did not test log x over the range $[1, 10^6]$. A casual glance at the algorithm used, however, is enough to convince us that the error has little to do with the size of x. (The first step expresses $x = 2^k m$, where $0 \leq m \leq 1$). In what follows we shall take M to be 10^{-10} to be safe.

<u>Lemma 2</u> If $\prod_{i=1}^{n} (1 + \varepsilon_i) = 1 + n\eta$, $n < \delta^{-1}$, and $|\varepsilon_i| \leq \delta$, then $|\eta| \leq \delta/(1-n\delta)$. Further, if $n \leq 100$, then $|\eta| \leq \delta/(1 - 100\delta)$.

<u>Proof</u> $\left|\prod_{i=1}^{n} (1 + \varepsilon_i) -1\right| \leq (1 + \delta)^n -1$

$$= n\delta + \binom{n}{2} \delta^2 + \ldots + \binom{n}{n} \delta^n$$

$$\leq n\delta + (n\delta)^2 + (n\delta)^3 + \ldots$$

$$= n\delta/(1 - n\delta) \leq n\delta/(1 - 100\delta).$$

<u>Remark</u> If $(1 + \varepsilon)^{-1} = 1 + \eta$, where $|\varepsilon| \leq \delta$, then $|\eta| \leq \delta/(1 - \delta)$.

In what follows we shall assume that $\delta = 10^{-12}$, $M = 10^{-10}$, $M' = M/(1 - 100M)$, and we shall adopt the convention that all $|\varepsilon_i| \leq \delta$,

and all $|\theta_i| \leqq M$. In particular, $|\varepsilon_i| \leqq \delta < M < M'$, and $|\Theta_i| \leqq M < M'$.

We now begin the error analysis for Programs One and Two which evaluate the sum $\sum_{Q*}\{\alpha(a,d) - .005\}/\sqrt{a}$, where $Q_1 \subset Q^* \subset Q$.

We first find the error in calculating each term

$$T = \{\alpha(a,d) - 0.005\}/\sqrt{a}.$$

We recall (1) that $\alpha(a,d) = \log a + \log (8\pi e^{-\gamma}) - (\frac{1}{2}) \log d$. It is a simple matter to determine from tables such as [5] that

$$2.64695576 < \log (8\pi e^{-\gamma}) = 3 \log 2 + \log\pi - \alpha < 2.64695577.$$

We defined C to be the lower value, and fed it into the machine in decimal; using C in place of $\log (8\pi e^{-\gamma})$ will never lead us to erroneously conclude that (3) holds.

Let $L = C - 0.005 - (\frac{1}{2}) \log d$.

$$\begin{aligned} \text{Then } f\ell(L) &= (1 + \varepsilon_1)\, f\ell(C - 0.005) - (1 + \varepsilon_2)\, f\ell((\tfrac{1}{2}) \log d) \\ &= (1 + \varepsilon_1)\, \{(1 + \varepsilon_3)\, f\ell(C) - (1 + \varepsilon_4)\, f\ell(0.005)\} \\ &\quad - (1 + \varepsilon_2)\,(1 + \varepsilon_5)\,(\tfrac{1}{2})\, f\ell(\log d) \\ &= (1 + \varepsilon_1)\, \{(1 + \varepsilon_3)\,(1 + \varepsilon_6)C - (1 + \varepsilon_4)\,(1 + \varepsilon_7)\, 0.005\} \\ &\quad - (1 + \varepsilon_2)\,(1 + \varepsilon_5)\,(\tfrac{1}{2})\,(1 + \theta_1) \log d \\ &= (1 + 3\eta_1)C - (1 + 3\eta_2)\, 0.005 - (1 + 3\eta_3)\,(\tfrac{1}{2}) \log d \end{aligned}$$

where $|\eta_i| < M(1 - 3M)^{-1}$ $(1 \leqq i \leqq 3)$. We used the fact that $f\ell(\frac{1}{2}) = \frac{1}{2}$ and that $f\ell(d) = d$ since d is an integer less than 10^6, and of course we used Lemma 2.

Our term T now is given by

$$T = (L + \log a)/\sqrt{a},$$

so that

$$fl(T) = \frac{(1 + \varepsilon_8)\ fl(L + \log a)}{fl(\sqrt{a})}$$

$$= \frac{(1 + \varepsilon_8)\ \{(1 + \varepsilon_9)\ fl(L) + (1 + \varepsilon_{10})\ fl(\log a)\}}{(1 + \theta_1)\ \sqrt{a}} \cdot$$

$$= \frac{(1 + \varepsilon_8)\ \{(1 + \varepsilon_9)\ fl(L) + (1 + \theta_2)(1 + \varepsilon_{10})\ \log a\}}{(1 + \theta_1)\ \sqrt{a}}$$

$$= \frac{(1 + 6\eta_4)C - (1 + 6\eta_5)\ 0.005 - (1 + 6\eta_6)(\frac{1}{2})\ \log d + (1 + 4\eta_7)\log a}{\sqrt{a}}$$

where $|\eta_i| < M'$. Therefore $|fl(T) - T|$

$$= |6\eta_4 C - 6\eta_5(0.005) - 3\eta_6 \log d + 4\eta_4 \log a| / \sqrt{a}$$

$$\leqq (18 + .03 + 3 \log 10^6 + 4 \log 10^3)\ M'$$

$$< 88\ M'.$$

Having bounded the absolute rounding error in T, the general term of the series $\sum_{Q*}(\alpha(a,d) - .005)/\sqrt{a}$, we now proceed to bound the error due to summing the terms of the series in floating point arithmetic. The number of terms is less than 4000 by (5). Let t_i $(1 \leqq i \leqq N)$ be an enumeration of the terms, $N < 4000$. Using a well-known result of Wilkinson -- see e.g. [2] -- we have

$$fl(\sum_{i=1}^{N} t_i) = \sum_{i=1}^{N} (1 + E_i)\ fl(t_i),$$

where $|E_i| < N(1 - N\delta)^{-1}\ \delta$. Since $fl(t_i) = t_i + u_i$, where $|u_i| \leqq 88\ M'$, we have

$$|fl\ (\sum_{i=1}^{N} t_i) - \sum_{i=1}^{N} t_i|$$

$$= |\sum_{i=1}^{N} (u_i + t_i\ E_i + E_i\ u_i)|$$

$$\leq 88\ NM' + N^2 (1 - N\delta)^{-1}\ \delta \max|t_i|$$

$$+ 88\ N^2 M' (1 - N\delta)^{-1}\ \delta.$$

It is easily seen that $|t_i| \leq 5$ for $d \leq 10^6$; the error is bounded by

$$88\ NM' + (5 + 88\ M')\ N^2\delta\ (1 - 4000\ \delta)^{-1}$$

$$< (5.1)\ N^2\delta + 88\ NM'.$$

If we now use the values $\delta = 10^{-12}$, $M = 10^{-10}$, and $N \leq 4000$, we see that the error is less than

$$(5.1)\ 16 \times 10^6 \times 10^{-12} + 90 \times 4000 \times 10^{-10}$$

$$< 1.2 \times 10^{-4}.$$

Hence any fundamental discriminant not written out onto tape by Program Two satisfies

$$\textstyle\sum_{Q*}\{\alpha(a,d) - .005\}/\sqrt{a} > 1/10 - 1.2 \times 10^{-4} > 0$$

and therefore satisfies (i) and (ii).

We now discuss the rounding error in Program Three, where the sum $\sum_Q\{\alpha(a,d) - H_o(a,b,c)\}/\sqrt{a}$ is evaluated for those fundamental discriminants $-d$ that failed Programs One and Two. This sum differs from the earlier sum only in that .005 is replaced by $H_o(a,b,c)$. Therefore we must estimate the errors in calculating $H_o(a,b,c)$ and k, which are given by (2).

We have $k \geq \sqrt{3}/2$, since $d = 4ac-b^2 \geq 4a^2-a^2 = 3a^2$, since by reduction $b^2 \leq a^2 \leq c^2$, and $k^2 = d/(4a^2) \geq (3/4)$.

In the following analysis, we still assume that $|\varepsilon_i| \leq \delta$ and $|\theta_i| \leq M$, and of course the ε_i and θ_i are unrelated to the previous ones. We begin with k.

$$f\ell(k) = f\ell(\sqrt{d}/(2a)) = \sqrt{d}(1 + \theta_1)(1 + \varepsilon_1)/(2a)$$

$$= \sqrt{d}(1 + \eta_1)/(2a) = k(1 + \eta_1),$$

where $|\eta_1| < 2M'$. We next consider $f\ell(\cos(\pi b/a)) = \cos(f\ell(\pi b/a)) + \theta_2$

$$= \cos(\pi b(1 + \eta_2)/a) + \theta_2$$

$$= \cos(\pi b/a)\cos(\eta_2\pi b/a) - \sin(\pi b/a)\sin(\eta_2\pi b/a) + \theta_2$$

where $|\eta_2| \leqq 2M'$. Now

$$|\cos(\eta_2\pi b/a) - 1| = 2\sin^2(\eta_2\pi b/(2a)) < (1/2)(\eta_2\pi b/a)^2 < (\pi^2/2)\eta_2^2 < 5(2M')^2 < M.$$

Also $|\sin(\eta_2\pi b/a)| \leqq \pi|\eta_2| \leqq 2\pi M'$. Hence, continuing, we have

$$f\ell(\cos(\pi b/a)) = (1 + \eta_3)\cos(\pi b/a) + \eta_4,$$

where $|\eta_3| \leqq M$, and $|\eta_4| \leqq 2\pi M' + |\theta_2| \leqq 8M$. Hence

$$f\ell(2\cos(\pi b/a)) = 2\{\cos(\pi b/a)(1 + \eta_3) + \eta_4\}(1 + \varepsilon_4)$$

$$= 2(1 + \eta_5)\cos(\pi b/a) + \eta_6,$$

where $|\eta_5| \leqq 2M'$ and $|\eta_6| \leqq 16M'$.

We next consider the exponential factor. We have

$$f\ell(\exp(-2\pi k)) = \exp(f\ell(-2\pi k))(1 + \theta_3) + \theta_4$$

$$= (1 + \theta_3)\exp\{-2\pi f\ell(k)(1 + \eta_7)\} + \theta_4$$

$$= (1 + \theta_3)\exp\{-2\pi k(1 + \eta_8)\} + \theta_4$$

$$= (1 + \theta_3)\exp(-2\pi k)\exp(-2\pi k\eta_8) + \theta_4,$$

where,

$|\eta_7| \leqq 2M'$ and $|\eta_8| \leqq 5M'$. Now $k = \sqrt{d}/(2a) \leqq (1/2)\sqrt{800{,}000}$ < 500, and hence $|\exp(-2\pi k\eta_8) - 1|$.

$$\leqq \frac{|2\pi k\,\eta_8|}{1 - |2\pi k\,\eta_8|}$$

$$\leqq \frac{2\pi \times 500 \times 5M'}{1 - 2\pi \times 500 \times 5M'}$$

$$< \frac{15800\, M'}{1-15800M'} < 15900\ M.$$

That is, $\exp\,(-2\pi k\eta_8) = 1 + E_1$ where $|E_1| < 15900\ M$.

Therefore

$$\begin{aligned} &f\ell(\exp(-2\pi k)) \\ &= (1 + E_1)(1 + \theta_3)\exp\,(-2\pi k) + \theta_4 \\ &= (1 + \eta_9)\exp\,(-2\pi k) + \theta_4, \end{aligned}$$

where $|\eta_9| < 15901\ M'$.

We now put

$$\beta = \frac{2\cos\,(\pi b/a)\exp\,(-2\pi k)}{\sqrt{k}}$$

Then, using different ε_i, we have

$$f\ell(\beta) = \frac{(1 + \varepsilon_1)(1 + \varepsilon_2) f\ell(2\cos(\pi b/a))\ f\ell(\exp\,(-2\pi k)).}{(1 + \theta_5)\sqrt{f\ell(k)}}$$

Now

$$\sqrt{f\ell(k)} = \sqrt{(1 + \eta_1)k} = (1 + E_2)\sqrt{k}\ ,\ \text{where } |E_2| < 2M'.$$

Hence

$$f\ell(\beta) = \frac{(1 + \varepsilon_1)(1 + \varepsilon_2)\{(1 + \eta_5)2\cos(\pi b/a)+\eta_6\}\{(1+\eta_9)\exp(-2\pi k) + \theta_4\}}{(1 + \theta_5)(1 + E_2)\sqrt{k}}$$

$$= (1 + \eta_{10})\beta + \eta_{11},\ \text{where } 1 + \eta_{10} = \frac{(1 + \eta_5)(1 + \eta_9)(1 + \varepsilon_1)(1 + \varepsilon_2)}{(1 + \theta_5)(1 + E_2)}$$

and

$$\eta_{11} = \frac{(1 + \varepsilon_1)(1 + \varepsilon_2)\{2(1 + \eta_5)\theta_4\cos(\pi b/a)+\eta_6(1 + \eta_9)\exp(-2\pi k)+\eta_6\theta_4\}}{(1 + \theta_5)(1 + E_2)\sqrt{k}}$$

Hence $|\eta_{10}| \leqq 16000M'$, and, since $\sqrt{k} \geqq (3/4)^{1/4} > 0.9$, we have

$$\eta_{11} \leqq \frac{2M' + 16M'\,(0.005) + 16MM'}{(1 - M)(1 - 2M')\,(.85)}$$

$$< \frac{2.1\ M'}{.8} < 3M.$$

We next calculate $H_o(a,b,c) = H = \beta + 0.04\ |\beta|$ using new ε_i.

$$f\ell(H) = f\ell(\beta + .04|\beta|) = f\ell(\beta)(1+\varepsilon_1) + .04|f\ell(\beta)|(1+\varepsilon_3)(1+\varepsilon_2)$$

$$= \beta(1+\eta_{10})(1+\varepsilon_1) + \eta_{11}(1+\varepsilon_1) + .04|\beta(1+\eta_{10}) + \eta_{11}|(1+\varepsilon_3)(1+\varepsilon_2).$$

Therefore,

$$f\ell(H) - H = \beta(\eta_{10}+\varepsilon_1(1+\eta_{10})) + \eta_{11}(1+\varepsilon_1) + .04|\beta|[(1+\eta_{10})(1+\varepsilon_3)(1+\varepsilon_2) - 1]$$

$$\pm\ .04|\eta_{11}|(1+\varepsilon_3)(1+\varepsilon_2).$$ Hence

$$|f\ell(H) - H| \leqq |\beta|\cdot|\eta_{12}| + |\eta_{13}| + 0.5 \times 16002M'\ |\beta| + .05\ \eta_{11}\ ,$$

where $|\eta_{12}| = |\eta_{10} + \varepsilon_1(1+\eta_{10})| \leqq |\eta_{10}| + 2|\varepsilon_1| \leqq 16002M'$, and $|\eta_{13}| \leqq 3M'$.

We require a bound on $|\beta|$. Now $|\beta| = |2k^{-1/2}\exp(-2\pi k)\ \cos(2b\pi/a)|$, and since $k \geq \sqrt{3}/2$, we have $|\beta| \leq |2(\sqrt{3}/2)^{-1/2}\exp(-2\pi(\sqrt{3}/2))| < 10^{-2}$. Therefore

$$|f\ell(H) - H| \leqq 10^{-2} \times 16002M' + 3M' + .05 \times 10^{-2} \times 16002M' + .05\ (3M')$$

$$\leqq 172M',$$

We now estimate the error for

$$T = \{C - H - (\tfrac{1}{2})\log d + \log a\}/\sqrt{a}\ .$$

The analysis is similar to that done previously except that H replaces .005.

Making the necessary changes, but using the same numbering on the η's, we have $f\ell(T) =$

$$\frac{(1 + 6\eta_4)C - (1 + 6\eta_5)f\ell(H) - (1 + 6\eta_6)(\frac{1}{2})\log d + (1 + 4\eta_7)\log a}{\sqrt{a}}$$

where $|\eta_i| < M'$.

Since $|f\ell(H) - H| \leqq 172M'$, and $|f\ell(H)| \leqq |H| + |f\ell(H) - H| \leqq 1.04|\beta| + 172M'$ $\leqq .0104 + 1.8 \times 10^{-8} < .0105$, we have

$$|f\ell(T) - T| =$$

$$\frac{|6\eta_4 C - \{f\ell(H) - H\} - 6\eta_5\ f\ell(H) - 3\eta_6 \log d + 4\eta_7 \log a\ |}{\sqrt{a}}$$

$\leq (18 + 172 + 6(0.0105) + 3 \log 10^6 + 4 \log 10^3)M'$

$< 260M'$.

In this particular program (Program Three), a count of the number of terms was recorded and found to be not more than 600. This is much less than the upper bound 4000 that was used on the previous estimate. Let t_i $(1 \leq i \leq N)$ be an enumeration of the terms of the sum $\sum_Q\{\alpha(a,d) - H_o(a,b,c)\}/\sqrt{a}$, $N \leq 600$.

The error analysis for the summation is similar to that for Programs One and Two.

$$f\ell(\textstyle\sum_{i=1}^N t_i) = f\ell\{\sum_{i=1}^N (t_i + u_i)\} = \sum_{i=1}^N (t_i + u_i)(1 + \gamma_i),$$

where $|u_i| < 260M'$ and $|\gamma_i| \leq N\delta(1 - N\delta)^{-1}$.

Therefore

$$\left|f\ell\left(\sum_{i=1}^N t_i\right) - \sum_{i=1}^N t_i\right|$$

$$= \left|\sum_{i=1}^N (u_i + t_i\,\gamma_i + u_i\,\gamma_i)\right|$$

$$\leq 260NM' + 5N^2\,\delta(1 - N\delta)^{-1} + 260\,M'\,N^2\delta(1 - N\delta)^{-1}$$

$$= 260NM' + (5 + 260M')\,N^2\delta\,(1 - N\delta)^{-1}$$

$$\leq 260 \times 600 \times 1.1\,M + 5.1 \times 3.6 \times 10^5\,\delta$$

If we take $M = 10^{-10}$ and $\delta = 10^{-12}$, the error is less than 2×10^{-5}. Thus the absolute error in calculating $\sum_Q\{\alpha(a,d) - H_o(a,b,c)\}/\sqrt{a}$ was less than 2×10^{-5}. When Program Three was run, the lowest value for the sum printed out was $.7 \times 10^{-4}$. Hence condition (ii) of Low's Hypothesis holds for the exceptional discriminants that failed Programs One and Two; hence (ii) holds for all d up to 800,000.

Now Program Three also verified (i) by calculating the sum $\sum_Q \alpha(a,d)/\sqrt{a}$. The error in this calculation is clearly less than the upper bound for the error in the calculation of $\sum_Q \{\alpha(a,d) - H_o(a,b,c)\}/\sqrt{a}$ and therefore less than 2×10^{-5}. Of the values for the sum $\sum_Q \alpha(a,d)/\sqrt{a}$ printed out by the computer, three were negative, (d = 115147, 357819, 636184; and a careful calculation showed them to be indeed negative), but none of the others were less than 1.9×10^{-4}. Hence condition (i) holds for all d less than 800,000, except for the three exceptions, where it definitely does not hold.

§4. Verification of (4) A second series of programs, analogous to Programs One and Two, combined, were run to verify (4). The first program tested whether $\sum_{Q_1} \{8/3 + \alpha^3(a,d)/6\}/\sqrt{a} \geqq 1/10$ (we use the fact that $b_3 > 8/3$), and the second program added on as many terms as necessary to get

$$\sum_{Q^*} \{8/3 + \alpha^3(a,d)/6\}/\sqrt{a} \geqq 1/10 \text{ for some } Q^*,\ Q_1 \subset Q^* \subset Q.$$

There were no failures to this program, and it was more economical of computer time than ones in the first series, because relatively few terms were needed beyond those from Q_1. We shall suppress the error analysis here which is very similar to the one for Programs One and Two of the first series.

§5. Summary Summarizing our results we have the following table.

If T means true, F means false, and U means unknown, we have the following tabulation for fundamental discriminants -d:

d	(i)	(ii)	(iii)
$1 < d < 115147$	T	T	T
$d = 115147$	F	T	T
$115147 < d < 357819$	T	T	T
$d = 357819$	F	T	T
$357819 < d < 636184$	T	T	T
$d = 636184$	F	T	T
$636184 < d \leqq 800{,}000$	T	T	T
$800{,}000 < d \leqq 1{,}200{,}000$	U	U	T

As in [1], we can use our results to estimate $L_{-d}(\frac{1}{2})$ and $L_{-d}(1)$ at the exceptional values. We get

$$L_{-115147}(\tfrac{1}{2}) \doteq .00007065$$

$$h(-115147) = 32$$

$$L_{-115147}(1) \doteq 0.29626$$

$$L_{-636184}(\tfrac{1}{2}) \doteq 0.017940$$

$$h(-636184) = 224$$

$$L_{-636184}(1) \doteq 0.882280$$

$$L_{-357819}(\tfrac{1}{2}) \doteq 0.0015555$$

$$h(-357819) = 112$$

$$L_{-357819}(1) \doteq 0.588215.$$

The two which Low had agree fairly well. He had

$$L_{-115147}(\tfrac{1}{2}) \doteq 0.000070675$$

and

$$L_{-636184}(\tfrac{1}{2}) \doteq 0.001794.$$

This leads me to think that possibly Low's omission of 357819 was a bookkeeping error.

REFERENCES

[1] M. E. Low, Real zeros of the Dedekind zeta function of an imaginary quadratic field, Acta Arithmetica, XIV(1969), 117-140.

[2] J. H. Wilkinson, Rounding Errors in Algebraic Processes, Prentice Hall, 1963.

[3] P. T. Bateman and E. Grosswald, On Epstein's zeta function, Acta Arithmetica, IX (1964), 365-373.

[4] K. E. Hillstrom, Performance Statistics for the Fortran IV (H) and PL/I (Version 5) Libraries in IBM OS/360 Release 18, Argonne National Laboratory Report ANL-7666, August 1970.

[5] Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Edited by Milton Abramowitz and Irene A. Stegun, Dover Publications, Inc., New York.

II. THE LATTICE TRIPLE PACKING OF SPHERES IN EUCLIDEAN SPACE

§1. Introduction

Let Λ be an n-dimensional lattice in n-dimensional Euclidean space E_n, such that, if open spheres of radius 1 are centered at the points of Λ, then no point of space is covered more than k times. That is, for any point X in E_n there do not exist distinct points $L_1, L_2, \ldots, L_{k+1}$ of Λ such that $|X-L_1|, \ldots, |X-L_{k+1}| < 1$. They we say that Λ provides a k-fold packing for spheres of radius 1. The terms single, double and triple are synonymous with k-fold for k=1,2 and 3.

Let $d(\Lambda)$ denote the determinant of Λ, and let $\Delta_k^{(n)}$ denote the lower bound of $d(\Lambda)$, taken over all lattices Λ that provide a k-fold packing for spheres of radius 1. (Thus $\Delta_1^{(n)}$ is the critical determinant of a sphere of radius 2.) It is well known and easy to see (e.g., divide one generator of the lattice by k) that $\Delta_k^{(n)} \leqq \Delta_1^{(n)}/k$.

It has been shown by Few [1] that $\Delta_2^{(2)} = (1/2)\Delta_1^{(2)}$, and Heppes [5] showed that $\Delta_k^{(2)} = \Delta_1^{(2)}/k$ if and only if $k \leqq 4$.

In [4] Few and Kanagasabapathy determined the exact value of $\Delta_2^{(3)}$, namely $3\sqrt{3/2}$, which is less than $\Delta_1^{(3)}/2 = 2\sqrt{2}$. By constructing particular lattices they also showed that $\Delta_2^{(n)} < \Delta_1^{(n)}/2$ for every $n \geqq 3$.

Few remarks in [2] that $\Delta_2^{(3)}$ is the only multiple packing constant known exactly in three dimensions or more, and in this note I shall prove that $\Delta_3^{(3)} \leqq 8\sqrt{38}/27 < \Delta_1^{(3)}/3 = 4\sqrt{2}/3$ and give evidence suggesting that $\Delta_3^{(3)} = 8\sqrt{38}/27$.

In fact, I prove:

Theorem 1 A certain lattice Λ_o of determinant $d_o = 8\sqrt{38}/27$ provides a

triple packing for the unit sphere S. Also Λ_o has generators P,Q,R with $|P| = 2/3$.

Theorem 2 Any lattice Λ having generators P',Q',R' with $|P'| \leqq 0.95$ providing a triple packing for S must have determinant $d(\Lambda) \geqq d_o$ with equality only when $\Lambda = \Lambda_o$. Hence Λ_o gives a local minimum of $d(\Lambda)$ for triple packing of unit spheres.

Remark There is extensive numerical evidence that $d(\Lambda)$ does not fall below d_o for any triple packing with S.

§2. An Economical Lattice Λ_o

Theorem 1 The best lattice triple packing for spheres in E^3 has determinant $d(\Lambda) \leqq 8\sqrt{38}/27 = \sqrt{2432/729} = 1.82649...$, since indeed the lattice Λ_o generated by P, Q and R where $P = (a,0,0) = (2/3,0,0)$, $Q = (h,b,0) = (1/3,\sqrt{3},0)$ and $R = (g,f,c)$, where $g = 1/3$, $f = (11\sqrt{3})/27$ and $c^2 = 3-f^2$, provides a triple packing for the unit sphere S.

Proof Convention: The letters λ, μ and ν will denote integers. $S(A,r)$ will be the open sphere of radius r centered at A; $S(A)$ will denote $S(A,1)$; thus $S = S(\text{origin})$. Suppose that the point $X = (x,y,z)$ is covered four times. Translating X by a lattice point, we may suppose that $X \varepsilon S$, and replacing X by $-X$ if necessary we may suppose $z \geqq 0$. The three other spheres covering X can be written $S(\lambda P+\mu Q+\nu R) = S+\lambda P+\mu Q+\nu R$ where $(\lambda,\mu,\nu) \neq (0,0,0)$. We must have $|\lambda P+\mu Q+\nu R| < 2$ since they must intersect S. Therefore

$$(*) \qquad (\lambda a+\mu h+\nu g) + (\mu b+\nu f)^2 + \nu^2 c^2 < 4$$

and $|\nu| < 2/c$. Since $c > 1$, we have $\nu \varepsilon\{-1,0,1\}$. Now ν cannot be -1, since otherwise $|X-(\lambda P+\mu Q-R)| \geqq |c+z| > 1$. Hence $\nu \varepsilon\{0,1\}$. From (*) we also get $|\mu b+\nu f| < 2$. Since $0 \leqq \nu \leqq 1$ and $0 \leqq f \leqq b/2$ and $b = \sqrt{3} > 4/3$ this gives $-2 < \mu < 2$, $\mu \varepsilon\{-1,0,1\}$. We divide the proof into two parts.

<u>Part 1</u> $y \geqq 0$. Then $\mu \in \{0,1\}$; in fact if $\mu = -1$, then for $X \in S(\lambda P+\mu Q+\nu R)$ we would have $|X-(\lambda P-Q+\nu R)|^2 \geqq (b+y-\nu f)^2 > \left(\frac{16}{27} b\right)^2 = \frac{256}{243}$, since $\nu \in \{0,1\}$. Also $(\mu,\nu) \neq (1,1)$, since $|\lambda P+Q+R|^2 \geqq b^2 + c^2 > 4$. Hence $(\mu,\nu) \in \{(0,0),(1,0),(0,1)\}$.

<u>Type 1 Spheres</u> Suppose $(\mu,\nu) = (0,0)$. Then $\lambda P+\mu Q+\nu R = \lambda P$, and $S(\lambda P) \cap S = \emptyset$ if $|\lambda| > 2$, since $|3P| = 3a = 2$. The $S(\lambda P)$ such that $0 < |\lambda| \leqq 2$ are called type 1 spheres.

<u>Type 2 Spheres</u> Suppose $(\mu,\nu) = (1,0)$. Then $\lambda P+\mu Q+\nu R = \lambda P+Q$, and $S(\lambda P+Q) \cap S = \emptyset$ if $\lambda \notin \{0,-1\}$, since then $|\lambda P+Q|^2 = b^2 + (\lambda a+h)^2 = 3+|2\lambda/3+1/3|^2 \geqq 4$. The $S(Q-P)$ and $S(Q)$ are called type 2 spheres.

<u>Type 3 Spheres</u> Suppose $(\mu,\nu) = (0,1)$. Then $\lambda P+\mu Q+\nu R = R+\lambda P$, $S(R+\lambda P) \cap S = \emptyset$ if $\lambda \notin \{0,-1\}$. To see this, observe that if $S(R+\lambda P) \cap S \neq \emptyset$ then $(\lambda a+g)^2 + f^2 + c^2 < 4$; since $f^2 + c^2 = 3$, $|2\lambda/3+1/3| < 1$ and $\lambda \in \{0,-1\}$. We call $S(R)$ and $S(R-P)$ type 3 spheres.

It follows from the discussion of type 2 and type 3 spheres that $S \cap S(\lambda P+E) = \emptyset$ if $\lambda \notin \{-1,0\}$ where $E \notin \{R,Q\}$. In particular,

$$\emptyset = S \cap S(E+P) = S(-P) \cap S(E) = S(-2P) \cap S(E-P),$$

$$\emptyset = S \cap S(E+2P) = S(-2P) \cap S(E),$$

$$\emptyset = S \cap S(E-2P) = S(P) \cap S(E-P) = S(2P) \cap S(E),$$

$$\emptyset = S \cap S(E-3P) = S(2P) \cap S(E-P).$$

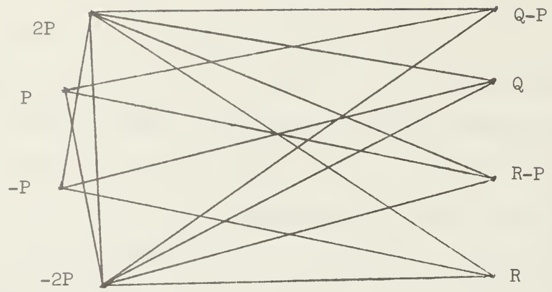
From the discussion of type 1 spheres $S \cap S(\lambda P) = \emptyset$ for $|\lambda| > 2$ so that

$$\emptyset = S \cap S(3P) = S(-P) \cap S(2P) = S(-2P) \cap S(P),$$

$$\emptyset = S \cap S(4P) = S(-2P) \cap S(2P).$$

We now draw a graph, G, where edges A and B are joined only if we know that $S(A) \cap S(B) = \emptyset$.

G:



We next observe that $\emptyset = S(Q+\lambda P) \cap S \cap S(R+\lambda' P+\nu Q)$. For $|Q+\lambda P| \geqq |Q| = \sqrt{b^2 + h^2} > \sqrt{3}$. Therefore the height (maximal value of the z coordinate of the closure) of $S(Q+\lambda P) \cap S$ is less than $\sqrt{1-3/4} = 1/2 < c-1$, since $c = 1.5\ldots > 3/2$. Since $R+\lambda' P$ has z component c, the above intersection is void. Hence we cannot have X simultaneously inside a sphere of type 2 and a sphere of type 3, so

$$X \in S(\lambda_1 P) \cap S(\lambda_2 P) \cap S(E+\lambda_3 P) \text{ with } 0 < |\lambda_1|, |\lambda_2| \leqq 2, -1 \leqq \lambda_3 \leqq 0$$

or

$$X \in S(\lambda_1 P) \cap S(E+\lambda_2 P) \cap S(E+\lambda_3 P) \text{ with } 0 < |\lambda_1| \leqq 2, -1 \leqq \lambda_2, \lambda_3 \leqq 0,$$

where $E \in \{R,Q\}$. Both of these contradict the graph G, and part 1 follows.

Part 2 We now suppose that $y < 0$. Recall that if $S \cap S(\lambda P+\mu Q+\nu R) \neq \emptyset$, then $-1 \leqq \mu \leqq 1$ and $0 \leqq \nu \leqq 1$. For those $S(\lambda P+\mu Q+\nu R)$ containing X we must have $-1 \leqq \mu \leqq 0$. For suppose that $\mu = 1$; since $X = (x,y,z)$, $y < 0$, we

would have $|\lambda P+ Q+\nu R-X| \geqq |b+\nu f-y| > b > 1$. The spheres $S(\lambda P+\mu Q+\nu R)$ containing X other than S may therefore be divided into four types, as follows:

Type 1 spheres, when $(\mu,\nu) = (0,0)$. As before the only spheres $S(\lambda P)$ intersecting S satisfy $0 < |\lambda| \leqq 2$, i.e., the type 1 spheres are $S(2P)$, $S(P)$, $S(-P)$ and $S(-2P)$.

Type 2 spheres, when $(\mu,\nu) = (-1,0)$. If $S(\lambda P+\mu Q+\nu R)$ is to intersect S we must have $4 > |\lambda P+\mu Q+\nu R|^2 = (\lambda a-h)^2 +b^2 = (2\lambda/3-1/3)^2 + 3$. Hence $0 \leqq \lambda \leqq 1$, i.e., the type two spheres are $S(-Q)$ and $S(P-Q)$.

Type 3 spheres, when $(\mu,\nu) = (0,1)$. As in part 1, $-1 \leqq \lambda \leqq 0$ if $S(\lambda P+\mu Q+\nu R)$ intersects S, i. e., the type 3 spheres are $S(R)$ and $S(R-P)$.

Type 4 spheres, when $(\mu,\nu) = (-1,1)$. If $S(\lambda P+\mu Q+\nu R)$ intersects S, then $(\lambda a+g-h)^2 + (f-b)^2 + c^2 < 4$, $4\lambda^2/9 < 4-(b-f)^2 - c^2 = 4/9$, $\lambda^2 < 1$, $\lambda = 0$, and $S(R-Q)$ is the only sphere of type 4.

As we did in part 1, we deduce several new disjoint pairs of spheres. From the discussion of type 4 spheres, $S \cap S(-Q+R+\lambda P) = \emptyset$ if $\lambda \neq 0$, so we have $S(\lambda P) \cap S(-Q+R) = \emptyset$ if $\lambda \neq 0$. From the type 2 spheres we have

$$\emptyset = S \cap S(-Q+2P) = S(-P) \cap S(-Q+P) = S(-2P) \cap S(-Q),$$

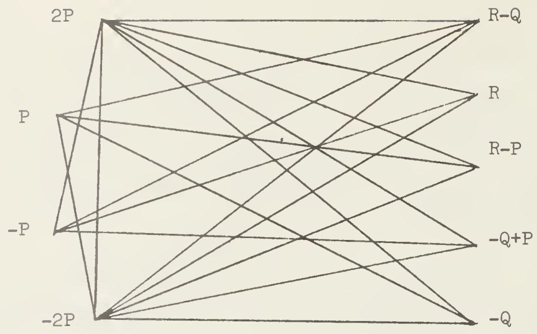
$$\emptyset = S \cap S(-Q+3P) = S(-2P) \cap S(-Q+P),$$

$$\emptyset = S \cap S(-Q-P) = S(P) \cap S(-Q) = S(2P) \cap S(-Q+P),$$

$$\emptyset = S \cap S(-Q-2P) = S(2P) \cap S(-Q).$$

If we combine these with some of the disjoint pairs that we already know from part 1 and draw a graph, G', in which A is joined to B only if we know $S(A) \cap S(B) = \emptyset$, we obtain

G':



In addition to these disjoint spheres, we observe that, for any λ and λ', $S(-Q+\lambda P) \cap S(R+\lambda' P) \cap S = \emptyset$, since $|-Q+\lambda P| \geq |Q|$, so that the height of $S(-Q+ P) \cap S$ is not greater than the height of $S \cap S(Q)$, which is less than $1/2 < c-1$, and c is the height of $R+\lambda' P$.

Also, for any λ, $\emptyset = S \cap S(Q+\lambda P) \cap S(R) = S \cap S(-Q-\lambda P) \cap S(-R) = S(R) \cap S(R-Q-\lambda P) \cap S$.

In particular,

(1) $\qquad S(R) \cap S(R-Q) \cap S = \emptyset.$

The following enumeration of possibilites shows that X cannot be contained in the necessary spheres, and theorem 1 follows:

Clearly types 2 and 3 cannot both occur by the fourth paragraph above, and two spheres of type 1 cannot occur with anything else by the graph G'. Again by the graph G', if two spheres of type 2 (or two spheres of type 3) occur, the remaining sphere cannot have type 1. Hence the only remaining possibilites are

(2) $\qquad X \in S(\lambda_1 P) \cap S(\lambda_2 P-Q) \cap S(R-Q), \quad 0 < |\lambda_1| \leqq 2,\ 0 \leqq \lambda_2 \leqq 1$

(3) $\qquad X \in S(\lambda_1 P) \cap S(\lambda_2 P+R) \cap S(R-Q), \quad 0 < |\lambda_1| \leqq 2,\ -1 \leqq \lambda_2 \leqq 0$

(4) $X \in S(P-Q) \cap S(-Q) \cap S(R-Q)$

or (5) $X \in S(R) \cap S(-P+R) \cap S(R-Q)$

Now (2) and (3) contradict G', and (1) excludes (5). To eliminate (4) we observe that, since

$$S(P) \cap S \cap S(R) \cap S(Q) = \emptyset ,$$

we must have

$$S(P-Q) \cap S(-Q) \cap S(R-Q) \cap S = \emptyset .$$

§3. The Lattice Λ_o is Locally Optimal

Remark 1 An arbitraty lattice Λ in E_3 has a basis P,Q,R where $|P| \leqq |Q| \leqq |R|$ are the successive minima of the unit sphere, $P = (a,0,0)$, $Q = (h,b,0)$, $R = (g,f,c)$, $a,b,c > 0$, $0 \leqq h \leqq a/2$, $0 \leqq f \leqq b/2$, and $-a/2 < g \leqq a/2$. Such a basis is said to be reduced in the sense of Gauss or simply reduced. For a proof, see [6], p. 163 et seq. "Seebers inequality".

Remark 2 If Λ has a reduced basis P,Q,R with $P = (a,0,0)$, $Q = (h,b,0)$, $R = (g,f,c)$ and if $d(\Lambda) \leqq d_o$, then $b^2 \leqq b_m^2$, where $b_m^2 = a^2/6+(2/3)\sqrt{a^4/16+3d_o^2/a^2}$.

Proof Using $|R^2| = g^2+f^2+c^2 \geqq |Q|^2 = b^2+h^2$, and the other inequalities of reduction, we have $d_o^2 \geqq d^2(\Lambda) = a^2b^2c^2 \geqq a^2b^2(b^2+h^2-g^2-f^2) \geqq a^2b^2(3b^2/4-a^2/4)$. Putting $t = b^2$, we get $3a^2t^2-a^4t-4d_o^2 \leqq 0$. Hence b^2 must lie between the roots $a^2/6-2/3\sqrt{a^4/16+3d_o^2/a^2}$ and $a^2/6+(2/3)\sqrt{a^4/16+3d_o^2/a^2}$ of the quadratic.

Let ρ^+ denote $\max\{0,\rho\}$.

Theorem 2 If (Λ,S) is a triple packing and $P = (a,0,0)$, $Q = (h,b,0)$ and $R = (g,f,c)$ gives a basis for Λ reduced in the sense of Gauss, and $a \leq 1$, then

(6) $$d(\Lambda) = abc \geqq ab\sqrt{\{4-(a+h)^2-((b^2-2ah-h^2)/(2b))^2\}^+} = f_1(a,b,h)$$

when $0 \leqq g \leqq h$, and

(7) $$d(\Lambda) = abc \geqq ab\sqrt{\{4-9a^2/4-(b^2+h^2-3ah)^2/(2b)^2\}^+} = f_2(a,b,h)$$

when $-a/2 \leqq g \leqq 0$ and when $h \leqq g \leqq a/2$. Furthermore $f_1(a,b,h) \geqq f_2(a,b,h)$, so that in fact

(8) $\quad d(\Lambda) \geqq f_2(a,b,h)$

in all cases. Also, if $d(\Lambda) \leqq d_o$ and $2/3 \leqq a \leqq 0.9508$, we have

(9) $\quad f_2^2(a,b,h) \geqq \min\{p(a),\ d_o^2 + 1/100\}$,

where $p(a) = 2a^6-11a^4+12a^2$, $p(2/3) = d_o^2$, and

(10) $\quad p(a) > d_o^2$ for $2/3 < a \leqq 0.9508$.

Hence $d(\Lambda) \geqq d_o$ for $2/3 \leqq a \leqq 0.9508$, and with equality only if $a = 2/3$.

<u>Proof</u> Suppose that (Λ,S) gives a triple packing and that P,Q,R form a reduced basis of Λ. From reduction, we have

(11) $\quad |P| \leqq |Q| \leqq |R|$,

$\quad 0 \leqq h \leqq a/2$,

$\quad 0 \leqq f \leqq b/2$, and

$\quad |g| \leqq a/2$.

We also have $a \geqq 2/3$, since otherwise the point $(1/2)P$ would be covered by $S(-P)$, S, $S(P)$ and $S(2P)$. Observe that the center of the parallelogram with vertices $P,Q,Q+P$ and the origin will be covered by the four spheres S, $S(P)$, $S(Q)$, and $S(Q+P)$ unless one of the diagonals $|Q+P|$, $|Q-P|$ is at least 2. Since $|Q+P| \geqq |Q-P|$ by (11), it follows that $4 \leqq |Q+P|^2 = b^2+(a+h)^2$; hence

(12) $\quad b^2 \geqq 4-(a+h)^2$, and $a \geqq 2/3$.

<u>Case 1</u> In this case we assume

(13) $\quad 0 \leqq g \leqq h$.

A consequence of (13) is that $|R-Q-P|$ is not less than $|R-Q+P|$. They cannot both be less than 2, since then the center of the parallelogram having vertices $R,Q,Q+P,R+P$ would be covered four times. Hence we have

(14) $\quad |R-Q-P| \geqq 2.$

Another consequence of (13) is that $|R+P|$ is not less than $|R-P|$. Considering the parallelogram with vertices P,R,R+P and the origin shows that

(15) $\quad |R+P| \geqq 2.$

With a view to proving (6) we imagine a,h,b to be fixed and find the point R = (g,f,c) having least non-negative c such that (13), (14) and (15) hold and also

(16) $\quad 0 \leqq f \leqq b/2.$

We are in fact looking for the lowest point X = (x,y,z) inside the rectangular prism given by

(17) $\quad 0 \leqq x \leqq h$

$0 \leqq y \leqq b/2$

$z \geqq 0,$

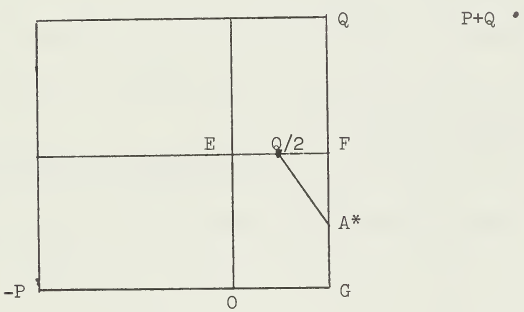
subject to the additional constraint

(18) $\quad |X-P-Q| \geqq 2$

$|X+P| \geqq 2.$

The problem is somewhat simplified by the fact that the centers P+Q, -P of the spheres lies on the plane z = 0, outside the prism.

Case 1.



If the right hand side of (6) is zero, there is nothing to prove. Let us suppose, therefore, that it is positive. We shall show that the point X^* that lies on the intersection of the boundary of $S(-P,2)$ and $S(P+Q,2)$ and the plane $x = h$ is the lowest point satisfying (17) and (18). We start by finding X^*. Let $X^* = (x^*,y^*,z^*)$. Let π be the radical plane of $S(-P,2)$ and $S(P+Q,2)$ (the plane obtained by subtracting the equations of the two spheres). Then π passes through $(1/2)Q = (h/2,b/2,0)$, which is halfway between the center of the two spheres, and has the equation $y-b/2 = (-(2a+h)/b)(x-h/2)$. Putting $x^* = h$, we obtain $y^* = b/2-(2ah+h^2)/(2b)$. We must show that $0 \leqq y^*$ so that (17) is satisfied. By (12) we have $b^2 \geqq 4-(a+h)^2 \geqq 2ah+h^2 + 1/2$, since $h \leqq a/2$; hence $2by^* = b^2-(2ah+h^2) > 0$ and (17) follows. We see that $z^* = \sqrt{4-(a+h)^2-((b^2-2ah-h^2)/(2b))^2} > 0$ by a previous assumption.

The first step in showing that X^* is optimal is to show that the bottom of the prism is covered by $S(-P,2)$ and $S(P+Q,2)$. This means that there is no X satisfying (17) and (18) with $z = 0$. Let $A^* = (x^*,y^*,0)$, $E = (0,b/2,0)$, $F = (h,b/2,0)$ and $G = (h,0,0)$. Then $2 = |P+Q-X^*| > |P+Q-A^*| \geqq |P+Q-F|$, and also $|P+Q-(1/2)Q| < 2$, since the spheres $S(-P,2)$ and $S(P+Q,2)$ intersect and $(1/2)Q$ is halfway between their centers. Hence the triangle with vertices $(1/2)Q$, A^*,F lies in the interior of $S(P+Q,2)$.

Similarly, $2 = |-P-X^*| > |-P-A^*| \geqq |-P-G| \geqq |-P|$ and $2 > |-P-(1/2)Q| \geqq |-P-E|$ so that the convex pentagon with vertices G, $A^*,(\frac{1}{2})Q$, E and the origin lies in the interior of $S(-P,2)$. Hence the bottom of the prism is covered.

We now let $X_1 = (x_1,y_1,z_1)$ be a lowest point satisfying (17) and (18). We know that X_1 exists, because the set of solutions is non-empty and closed.

The point X_1 must be on the boundary of $S(-P,2)$ or $S(P+Q,2)$ since otherwise it could be lowered and still satisfy (18).

Let us suppose first that X_1 is on the boundary of $S(-P,2)$. We shall deduce that X_1 is on the boundary of $S(P+Q,2)$. Suppose not. Then (x_1,y_1) must be the point satisfying (17) that is farthest from $-P$, namely $(h,b/2)$. But then the point $X_1 = (h,b/2,\sqrt{4-(h+a)^2-b^2/4})$ is easily seen to be inside $S(P+Q,2)$, contrary to (18).

Suppose that X_1 is not on the boundary of $S(-P,2)$. Then $|X_1-P-Q| = 2$, and $X_1 = (0,0,\sqrt{4-(a+h)^2-b^2})$ lies inside $S(-P,2)$ contrary to (18).

Hence X_1 lies on the arc of the intersection of $S(-P,2)$ and $S(P+Q,2)$ with $z \geqq 0$. The highest point of the arc is the point directly above $(1/2)Q$, which is on the boundary of the prism, and the lowest point in the prism is X^*, where the arc cuts the $x = h$ plane. Hence $X_1 = X^*$ and (6) follows.

<u>Case 2</u> Assume

(19) $\quad -a/2 \leqq g \leqq 0.$

Since the center of the parallelogram with vertices $R,Q-P,R+P,Q$ must not be covered four times, we know that one of its two diagonals $|R-Q|$, $|R-Q+2P|$ must be at least 2. Assuming Gauss reduction, we always have $|R-Q+2P| \geqq |R-Q|$, since the vectors $R-Q+2P$ and $R-Q$ differ only in the first component, and $|g-h+2a| \geqq 2a-|g|-h \geqq a \geqq |g-h|$. Hence $|R-Q+2P| \geqq 2$.

As in case 1, one of $|R+P|$, $|R-P|$ must be at least 2, and from (19) we know that $|R-P| \geqq |R+P|$ so that $|R-P| \geqq 2$.

In a manner similar to case 1, we are looking for the lowest point $X = (x,y,z)$ inside the rectangular prism given by

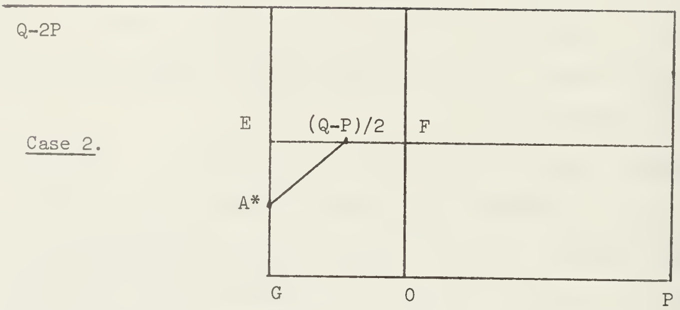
(20) $\quad -a/2 \leqq x \leqq 0$

$\quad 0 \leqq y \leqq b/2, \; z \geqq 0$

such that

(21) $|X-P| \geqq 2$, and $|X-Q+2P| \geqq 2$.

Case 2.



The procedure is the same as in case 1. We may suppose that the right hand side of (7) is positive. Let $X^* = (x^*,y^*,z^*)$ be the point on the two spheres and the plane $x = -a/2$, with $z^* \geqq 0$.

The radical plane π of the two spheres passes through $(1/2)(Q-P)$. The equation of π is $2(h-3a)x+2by = b^2+(h-3a)(h-a)$. Putting $x = x^* = -a/2$ yields $y^* = (b^2+h^2-3ah)/(2b)$. We must show that $0 \leqq y \leqq b/2$. Now $b^2 \geqq 4-(a+h)^2$ for a triple packing; hence $2by^* = b^2+h^2-3ah \geqq 4-a^2-5ah \geqq 1/2$. On the other hand $h^2 \leqq ah/2 \leqq 3ah$, $2by^* = b^2+h^2-3ah \leqq b^2$, $y^* \leqq b/2$. We see that $z^* = \sqrt{4-(3a/2)^2-((b^2+h^2-3ah)/(2b))^2} > 0$ by assumption.

We now show that the bottom of the prism is covered by the two spheres $S(P,2)$ and $S(Q-2P,2)$. Let $A^* = (x^*,y^*,0)$, $E = (-a/2,b/2,0)$, $F = (0,b/2,0)$ and $G = (-a/2,0,0)$. Then $2 = |Q-2P-A| > |Q-2P-A^*| \geqq |Q-2P-E|$, and also $|(Q-2P)-(1/2)(Q-P)| < 2$ since the spheres $S(P,2)$ and $S(Q-P,2)$ intersect and $(1/2)(Q-P)$ is halfway between their centers. Hence the triangle with vertices

$(1/2)(Q-P)$, E, A^* lies in the interior of $S(Q-2P,2)$. Similarly $2 > |P-(1/2)(Q-P)| \geqq |P-F| \geqq |P|$ and $2 > |P-A^*| \geqq |P-G|$, so that the convex pentagon with vertices G, A^*, $(1/2)(Q-P)$, F, and the origin lies in the interior of $S(P,2)$. Hence the bottom of the prism is covered.

We now let $X_1 = (x_1, y_1, z_1)$ be a lowest point satisfying (20) and (21); X_1 exists because the set of points satisfying (20) and (21) is closed and non-empty. Since $z_1 > 0$, X_1 must be on the boundary of one of the spheres. We suppose first that $|X_1-P| = 2$, and $|X_1-Q+2P| > 2$. Then (x_1, y_1) must be as far from $(a,0)$ as possible still satisfying (20). Hence $X_1 = (-a/2, b/2, \sqrt{4-b^2/4-(3a/2)^2})$, and calculation shows that $|X_1-Q+2P| < 2$, which is a contradiction.

Suppose now that $|X_1-P| > 2$, so that $|X_1-Q+2P| = 2$. Then (x_1, y_1) must be as far as possible from $(h-2a, b)$ and still satisfy (20). Now $-2a \leqq h-2a \leqq -(3/2)a < -a/2$. Hence $X_1 = (0,0,\sqrt{4-(h-2a)^2-b^2})$. Hence $|X_1-P|^2 = a^2+4-(h-2a)^2-b^2 < a^2+4-9a^2/4 < 4$, which is a contradiction.

Hence X_1 must be on the boundary of both spheres. In a manner similar to case 1, X_1 lies on a circular arc whose highest point $(1/2)(Q-P)$ is on the boundary of the prism and whose lowest point inside the prism is X^*. Hence $X_1 = X^*$ and (7) is proved for this case.

<u>Case</u> 3 Assume

$$h \leqq g \leqq a/2. \tag{22}$$

The vectors $R-Q+P$, $R-Q-P$ differ only in the first component and, by (22), $|g-h+a| \geqq |g-h-a|$. Since the center of the parallelogram with vertices $P, R-Q, R-Q+P$ and the origin must not be covered four times, we conclude that $|R-Q+P| \geqq 2$. On the other hand, as in case 1, $|R+P| \geqq 2$.

We are seeking the lowest point X_1 of the prism

(23) $h \leq x \leq a/2$,

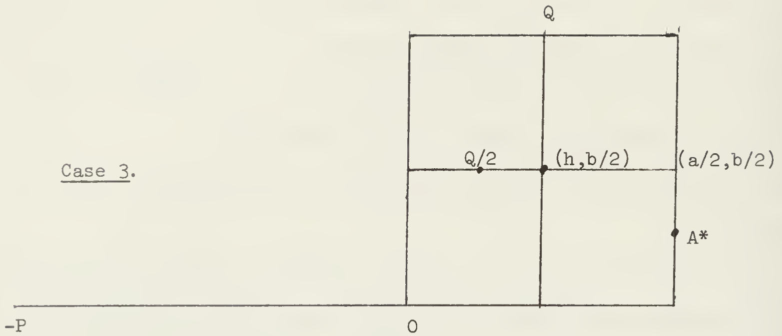
$0 \leq y \leq b/2$,

$z \geq 0$,

such that

(24) $|X+P| \geq 2$, $|X-Q+P| \geq 2$.

Case 3.



If the right hand side of (7) is zero, there is nothing to prove. Let us suppose that it is positive. Let X be the point on the two spheres and the plane $x = a/2$. To find $X = (x,y,z)$ we solve

$$9a^2/4+y^2+z^2 = 4$$

$$(3a/2-h)^2+(y-b)^2+z^2 = 4$$

and obtain $x^* = a/2$, $y^* = (b^2+h^2-3ah)/(2b)$, and $z^* = \sqrt{4-9a^2/4-(y^*)^2}$. This is the same y^* that appears in case 2, so $0 \leq y^* \leq b/2$ and X^* satisfies (23) and (24).

Now let X_1 be a lowest point satisfying (23) and (24), and we shall show that $X_1 = X^*$.

We must show that the bottom of the prism is covered. Since $x = a/2$ maximizes the horizontal distance from both -P and Q-P, it is sufficient to show that the line segment $\{(a/2,t,0) : 0 \leq t \leq b/2\}$ is covered. The spheres

$S(-P,2)$ and $S(Q-P,2)$ intersect the plane $x = a/2$ in circles which intersect at $(a/2,y^*,z^*)$. Since $z^* > 0$, the segment is covered.

For a lowest point X_1 we must have $x_1 = a/2$, Since the spheres intersect the $x = a/2$ plane in circles whose centers have y components 0 and b respectively, it is clear that $X_1 = X^*$ and inequality (7) holds. This finishes case 3.

Hence $d(\Lambda) \geqq \min\{f_1,f_2\}$. We observe immediately that $f_1 \geqq f_2$. This would follow if $\psi(h) \geqq 0$ where $\psi(h) = 9a^2/4+(b/2-(3a-h)h/(2b))^2-(h+a)^2-\{(b/2-(h/2)(h+2a)/b)\}^2$. Now $\psi(0) = 9a^2/4-a^2 =(5/4)a^2 > 0$ and $\psi(a/2) = 0$. Differentiating, we have $\psi'(h) = -\{(5a)/(2b^2)\}(b^2+3h^2-ah) < 0$ since $b^2 \geqq a^2-h^2 \geqq (3/4)a^2$.

We now prove

(9) $\qquad f_2^2 \geqq \min\{d_o^2 + 1/100,\ 2a^6-11a^4+12a^2\}$

under the hypothesis that $d(\Lambda) \leqq d_o$ and $2/3 \leqq a \leqq 0.9508$. Write $F(a,h,b) = f_2^2 = a^2b^2\{4-9a^2/4-(b/2+h(h-3a)/(2b))^2\}$ and put $t = b^2$. Clearly $\partial^2F/\partial t^2 = -a^2/2 < 0$. Hence $F(a,h,b) \geqq \min\{\phi(a,h),\psi(a,h)\}$, where $\phi(a,h) = F(a,h,\sqrt{4-(a+h)^2})$, and $\psi(a,h) = F(a,h,b_m)$, where b_m was defined in remark 2. Now $\phi(a,h) = a^2\{4-(a+h)^2\}\{4-9a^2/4-(1/4)(4-(a+h)^2)-(1/2)h(h-3a)\}-(a^2/4)h^2(h-3a)^2$ and calculation shows that $\partial^2\phi/\partial h^2= -8a^2-8a^4 < 0$. Since $\phi(a,0) = \phi(a,a/2) = p(a)$ where $p(x) = 2x^6-11x^4+12x^2$, we have $\phi \geqq p(a)$.

To complete the proof of (9), we will show that $\psi(a,h) \geqq d_o + 1/100$. Now $\psi(a,h) = F(a,h,b_m) = a^2b_m^2\{4-9a^2/4-(b_m/2+h(h-3a)/(2b_m))^2\}$ and calculation shows that $\frac{\partial^2\psi}{\partial h^2} =-a^2b_m^2+3a^2h^2-9a^3h+9a^4/4$. We shall show that $\frac{\partial^2\psi}{\partial h^2} < 0$, so that $\psi(a,h)$ is a concave function of h.

We digress for a moment to show that $h \geqq h_m(a)$, where $h_m = \sqrt{4-b_m^2} -a$.

To see this, recall that for triple packing we must have $b^2 \geqq 4-(a+h)^2$, whereas $b^2 \leqq b_m^2$, since $d(\Lambda) \leqq d_o$. The juxtaposition $b_m^2 \geqq 4-(a+h)^2$ yields $h \geqq h_m$. Putting $h = a/2$ we see that $h_m \leqq a/2$. It is conceivable that h_m is negative, even though h never is. For what follows, it is useful to know that $h_m \geqq -a/2$. To see this, note that

$$b_m^2 = a^2/6+(2/3)\sqrt{a^4/16 + 3d_o^2/a^2}$$

$< 1/6+(2/3)\sqrt{1/16 + 27} < 3.64$, since $2/3 \leqq a \leqq 1$ and $3 < d_o^2 < 4$, so that $4-b_m^2 > 1/4 \geqq a^2/4$, $h_m + a = \sqrt{4-b_m^2} > a/2$, and $h_m > -a/2$.

We now return to showing that $\psi(a,h)$ is a concave function of h for $h_m \leqq h \leqq a/2$. Since $\frac{\partial^3\psi}{\partial h^3} = 6a^2h-9a^3 \leqq 3a^3-9a^3 = -6a^3 < 0$, it is enough to show that $f(a) = \frac{1}{a^2}\left.\frac{\partial^2\psi}{\partial h^2}\right|_{h=h_m} < 0$.

Since f(a) is a rather complicated function of the single variable a, we shall simply find an upper bound for its derivative as a function of a and use a computer to evaluate it on a fine grid.

Let $F(x,y,z) = -z+3y^2-9xy+9x^2/4$ so that $F(a,h_m,b_m^2) = f(a)$. Then

$$|f'(a)| \leqq \left|\frac{\partial F}{\partial x}\right|_o + \left|\frac{\partial F}{\partial y}\right|_o \left|\frac{dh_m}{da}\right| + \left|\frac{\partial F}{\partial z}\right|_o \left|\frac{d}{da} b_m^2\right|,$$

where the subscript o indicates that the partial derivatives are evaluated at $(x,y,z) = (a,h_m,b_m^2)$.

Then

$$\left|\frac{\partial F}{\partial x}\right|_o = |-9h_m+9a/2| \leqq 9a \leqq 9,\quad \left|\frac{\partial F}{\partial y}\right|_o \leqq |6h_m| + 9a \leqq 12, \text{ and } \left|\frac{\partial F}{\partial z}\right|_o = 1.$$

Let $u = a^2$, $g(u) = b_m^2 = \frac{u}{6} +(2/3)\sqrt{u^2/16 + 3d_o^2/u}$. Then

$g'(u) = 1/6 +(1/3)(u/8-3d_o^2/u^2)/\sqrt{u^2/16+3d_o^2/u}$,

and

$|g'(u)| \leq 1/6 + (1/3)(1/8+3 \times 4/(4/9)^2)/\sqrt{(4/9)^2/16+9} = 1/6 +$

$(1/3)(1/8+243/4)/\sqrt{9} < 7.$ Therefore

$\left|\frac{d}{da} g(a^2)\right| = |2ag'(a^2)| < 14.$ That is, $\left|\frac{db_m^2}{da}\right| < 14$.

Finally,

$$h_m = \sqrt{4-b_m^2} - a, \quad \frac{dh_m}{da} = -1-(1/2)\frac{db_m^2}{da} \Big/ \sqrt{4-b_m^2},$$

and

$$\left|\frac{dh_m}{da}\right| \leq 1 + 7\{4-b_m^2\}^{-1/2} < 1 + 7\{0.36\}^{-1/2} < 13, \text{ since } b_m^2 < 3.64. \text{ Hence}$$

$$|f'(a)| \leq 9 + 12 \times 13 + 1 \times 14 = 179.$$

Using a computer we verified that (allowing for roundoff error) $f(a_i) < -0.2$ at the points $2/3 = a_o < a_1 < \ldots < a_n = 1$, where $n = 500$, and $|a_{i-1} - a_i| < \frac{1}{1200}$ for $1 \leq i \leq n$.

Let a be an arbitrary number in the interval $[2/3,1]$. Then $a \varepsilon [a_{i-1}, a_i]$ for some i, and therefore

$$f(a) = f(a_{i-1}) + \int_{a_{i-1}}^{a} f'(t)dt$$

$$\leq -0.2 + \frac{179}{1200} < -0.05.$$

Hence $\psi(a,h)$ is a concave function of h as claimed, so

$$\psi(a,h) \geq \min\{\psi(a,h_m), \psi(a,a/2)\}.$$

The functions $\psi(a,h_m)$ and $\psi(a,a/2)$ are also rather complicated functions of the single variable a, and they are both above $d_o^2 + 1/100$. We shall simply find an upper bound for their derivatives as functions of a^2 or a and use a computer to evaluate them on a fine mesh.

Let $u = a^2$ and $f(u) = \psi(a,a/2)$. Then $f(u) = ug(u)\Omega(u,g(u))-25u^3/64$ where $g(u) = b_m^2 = u/6+(2/3)\sqrt{u^2/16+3d_0^2/u}$, and $\Omega(u,v) = 4-(9/4)u-v/4+(5/8)u$. Then $f'(u) = g(u)\Omega(u,g(u))+ug'(u)\Omega(u,g(u))+ug(u)\{\partial\Omega/\partial u+g'(u)\partial\Omega/\partial v\}-75u^2/64$.

To estimate $|f'(u)|$, we must estimate g,Ω, and their derivatives. We have $|g(u)| \leqq 3.64 < 4$ from before and $|g(u)| \geqq u/3$, trivially. Hence $|\Omega(u,g(u)| \leqq 4+13u/8+|g(u)/4| \leqq 4+13/8+1 \leqq 7$. We also have $|g'(u)| \leqq 7$ from before. On the other hand $|\partial\Omega/\partial u| = |13/8| < 2$ and $|\partial\Omega/\partial v| = 1/4$. Putting the estimates together, $|f'(u)| < 4 \times 7 + 7 \times 7 + 4 \times 2 + 5 + 2 = 92$.

We will show that $f(u) \geqq d_0^2 + 1/100$ for $4/9 \leqq u \leqq 1$. Let us suppose that a computing machine has verified that $f(u_i) \geqq d_0^2 + 1/50 + \varepsilon$ for $4/9 = u_0 < u_1 < \cdots < u_n = 1$, where $|u_{i-1} - u_i| < (50 \times 92)^{-1}$. It then follows that $f(u) \geqq d_0^2 + \varepsilon$ for $4/9 \leqq u \leqq 1$. For $u\varepsilon[u_{i-1},u_i]$ for some i, and $|f(u)| \geqq |f(u_{i-1})| - |\int_{u_{i-1}}^{u} f'(t)dt| \geqq d_0^2 + 1/50 + \varepsilon - 92|u_{i-1} - u_i| \geqq d_0 + \varepsilon$.

A computing machine was programmed to find the minimum value of $f(u_i)$ for $1 \leqq i \leqq 8251$, where $u_i = 4/9 + (i-1)/14850$, and the answer was 3.51822.

Had there been no round-off error, we could say that $f(u_i) \geqq d_0^2 + 1/6$. It is certainly safe to say that $f(u) \geqq d_0^2 + 1/50 + 1/100$. Hence $\psi(a,a/2) = f(a^2) > d_0^2 + 1/100$ for $2/3 \leqq a \leqq 1$.

We shall use the same method for $\psi(a,h_m)$. Let us rename $f(a) = \psi(a,h_m) = a^2b_m^2\{4-(9/4)a^2- b_m^2/4 - h_m(h_m-3a)/2\} - a^2h_m^2(h_m-3a)^2/4$. It is unfortunate that $\psi(a,h_m)$ cannot be written simply as a function of a^2; all our functions are now functions of a. Let $g_1(a) = b_m^2 = g(a^2)$, and let $h(a) = h_m$. Put $\Phi(x,y,z) = 4 - (9/4)x^2-y/4 - z(z-3x)/2$ and $\Theta(x,z) = -x^2z^2(z-3x)^2/4$. Then $f(a) = a^2g_1(a)\Phi(a,g_1(a),h(a)) - \Theta(a,h(a))$, and $f'(a) =$

$2ag_1(a)\Phi(a,g_1(a),h(a)) + a^2g_1'(a)\Phi(a,g_1(a),h(a)) + a^2g_1(a)\{\frac{\partial\Phi}{\partial x} +$ $(\partial\Phi/\partial y)g_1'(a) + (\partial\Phi/\partial z)h'(a)\} - \partial\theta/\partial x - (\partial\theta/\partial z)h'(a)$. Using some of the estimates from before and making some new ones, we see that $|g_1(a)| = |g(a^2)| < 4$, $|h(a)| = |\sqrt{4-b_m^2} - a| \leqq \frac{1}{2}$, $|\Phi| \leqq 4 + 9/4 + 1 + h^2(a)/2 + 3|h(a)|/2 \leqq 1/8 + 28/4 < 8$, $|g_1'(a)| < 14$, $\partial\Phi/\partial x = 9x/2 + 3z/2$, $|\partial\Phi/\partial x| \leqq 9/2 + (3/2) \times \frac{1}{2} < 6$, $|\partial\Phi/\partial y| = 1/4$, $|\partial\Phi/\partial z| = |-z + 3x/2| \leqq \frac{1}{2} + 3/2 = 2$, and $|h'(a)| < 13$. By calculation, $\partial\theta/\partial x = xz^2(z^2-9xz+18x^2)/2$. Taking the maximum of the positive and negative parts, $|\partial\theta/\partial x| \leqq h_m^2/2 \times \max\{h_m^2+18a^2, 9h_m\}$ $\leqq 1/8 \max\{1/4+13, 9/2\} < 3$. Similarly, $\partial\theta/\partial z = -x^2z(z^2-6zx+9x^2+z^2-3zx)/2$, and $|\partial\theta/\partial z| \leqq (h_m/2)\max\{2h_m^2+9, 9h_m\} < 3$. Putting these estimates together, $|f'(a)| \leqq 2 \times 4 \times 8 + 14 \times 13 + 4(6+14/4+2\times 13) + 3 + 3 \times 13 = 430$. To show that $f(a) \geqq d_o^2 + \varepsilon$, therefore, it is enough to show that $f(a_i) \geqq d_o^2 + 1/50 + \varepsilon$ for $2/3 = a_o < \cdots < a_n = 1$ where $|a_{i-1} - a_i| < 1/25{,}000$.

A computing machine was programmed to find the minimum value of $f(a_i)$ for $1 \leqq i \leqq 100{,}001$ where $a_i = 2/3 + i/300{,}000$, and the answer was 3.40344 . . .

Had there been no round-off error we could say that $f(a_i) \geqq d_o^2 + 1/15$. It is certainly safe to say that $f(a_i) \geqq d_o + 1/50 + 1/100$ so that $\psi(a,h_m) = f(a) \geqq d_o^2 + 1/100$ for $2/3 \leqq a \leqq 1$. Therefore $\psi(a,h) \geqq \min\{\psi(a,a/2), \psi(a,h_m)\} \geqq d_o^2 + 1/100$ and $f_2^2 \geqq \min\{\psi(a,h),\phi(a,h)\} \geqq \min\{d_o^2 + 1/100, p(a)\}$ as claimed. Thus (9) is proved.

We now prove (10). Let $f(t) = 2t^3 - 11t^2 + 12t - d_o^2$. Then $f'(t) = 6t^2 - 22t + 12$ and $f''(t) = 12t - 22 < 0$ for $0 \leqq t \leqq 1$. Hence $f(t)$ is a concave function and has at most two zeroes in the range $[0,1]$. In fact, $f(4/9) = 0$, and $f(\alpha^2) = 0$ where $\alpha^2 = 0.90402\ldots$ and $f(t) > 0$ for $4/9 < t < \alpha^2$. Since $f(a^2) = p(a) - d_o^2$, we conclude that $p(a) > d_o^2$ for $2/3 < a < \alpha$ =

0.950802... and (10) is proved.

It now follows from (8), (9) and (10) that $d(\Lambda) \geqq d_0$ for $2/3 \leqq a \leqq 0.9508$ with equality only when $a = 2/3$.

REFERENCES

[1] L. Few, Ph.D. Thesis, London, (1953).

[2] L. Few, Multiple packing of spheres: a survey, Proc. Coll. Convexity, Copenhagen, 1965 (1967), 88-93.

[3] L. Few, The double packing of spheres, Journal London Math. Society, 28 (1953), 297-304.

[4] L. Few and P. Kanagasabapathy, The double packing of spheres, Journal London Math. Society, (series 1), 44 (1969), 141-146.

[5] A. Heppes, Mehrfache gitterformige kreislagerungen in der ebene, Acta Math. Acad. Sci. Hung., 10/1-2 (1959), 141-148.

[6] L. E. Dickson, Studies in the Theory of Numbers, University of Chicago Press.

III. SOME EXTREMAL PROBLEMS IN GEOMETRY

§1. Introduction

Let there be given n points $X_1,\ldots,X_n$ in k-dimensional Euclidean space E_k. Denote by $d(X_i,X_j)$ the distance between X_i and X_j. Let $A(X_1,\ldots,X_n)$ be the number of distinct values of $d(X_i,X_j)$, $1 \leqq i \leqq j \leqq n$. Put $f_k(n) = \min A(X_1,\ldots,X_n)$, where the minimum is taken over all possible choices of distinct $X_1,\ldots,X_n$. Denote by $g_k(n)$ the maximum number of solutions of $d(X_i,X_j) = \alpha$, $1 \leqq i \leqq j \leqq n$, where the maximum is to be taken over all possible choices of α and n distinct points $X_1,\ldots,X_n$. The estimation of $f_k(n)$ and $g_k(n)$ are difficult problems even for k=2. It is known that [1,2]

(1) $$cn^{2/3} < f_2(n) < Cn/\sqrt{\log n} \qquad \text{and}$$

(2) $$n**(1+c/\text{loglog } n) < g_2(n) < Cn^{3/2},$$

where c and C are positive absolute constants and a**b denotes a^b.

If $k \geqq 4$ the study of $g_k(n)$ becomes somewhat simpler [3].

A. Oppenheim posed the problem of investigating the number of triangles chosen from n points in the plane which have the same non-zero area. This question and its generalization were the first investigated in [6]. In this note I support some claims made in [6].

§2. Notations

Let $n \geqq 3$, $X_1,\ldots,X_n$ be n points in k-dimensional space E_k, and let $\Delta>0$.

We define $g_k^{(2)}(n;\, X_1,\ldots,X_n;\, \Delta)$ to be the number of triangles of the form $X_iX_jX_k$ having area Δ. We let

$$g_k^{(2)}(n;\, X_1,\ldots,X_n) = \max_{\Delta} g_k^{(2)}(n;\, X_1,\ldots,X_n;\, \Delta)$$

and
$$g_k^{(2)}(n) = \max_{X_1,\ldots,X_n} g_k^{(2)}(n;\, X_1,\ldots,X_n).$$

Let P be a fixed point and define $G_k^{(2)}(n; X_1,\dots,X_n; \Delta)$ to be the number of triangles of the form PX_iX_j having area Δ. We let

$$G_k^{(2)}(n) = \operatorname*{Max}_{X_1,\dots X_n,\ \Delta>0} G_k^{(2)}(n; X_1,\dots X_n; \Delta).$$

Clearly $g_{k-1}^{(2)}(n) \leqq g_k^{(2)}(n) \leqq nG_k^{(2)}(n-1) \leqq nG_k^{(2)}(n)$. We see that $g_k^{(2)}(n)$ is analogous to $g_k(n)$.

§3. The Article of Erdős and Purdy

It was shown [6] that

(3) $$cn^2 \log\log n \leqq g_2^{(2)}(n) \leqq nG_2^{(2)}(n) \leqq 4n^{5/2},$$

where c is a positive absolute constant, and

(4) $$g_2^{(2)}(n) \leqq g_3^{(2)}(n) \leqq nG_3^{(2)}(n) \leqq c\, n^{3-1/3}.$$

A simple example, which I shall give in §4 shows that $G_4^{(2)}(n) \geqq cn^2$ and $g_6^{(2)}(n) \geqq cn^3$. It is therefore worth asking whether $g_4^{(2)}(n)$ and $g_5^{(2)}(n)$ are $o(n^3)$. The object of this note is to support the claim made in [6] that in fact $g_4^{(2)}(n) \leqq g_5^{(2)}(n) \leqq cn^{3-\varepsilon}$ for some $\varepsilon > 0$.

§4. The Example of Linz Generalized

We first give the example that shows that $G_4^{(2)}(n) \geqq cn^2$. Let $n \geqq 2$ be given. Let $n = 2m+r$ where $0 \leqq r < 2$. Choose a co-ordinate system in E_4 and put $X_i = (a_i, b_i, 0, 0)$, for $1 \leqq i \leqq m$ and $Y_i = (0, 0, a_i, b_i)$, for $1 \leqq i \leqq m+r$, where (a_i, b_i) are $m+r$ distinct real solutions of $a^2 + b^2 = 1$. Then the $m(m+r)$ triangles OX_iY_j are all congruent to the triangles with sides 1, 1, $\sqrt{2}$ and therefore have the same (positive) area. Hence $G_4^{(2)}(n) \geqq m(m+r) \geqq n^2/4 - 1/4 \geqq cn^2$. By choosing the a_i, b_i so that some of the triangles OY_iY_j and OX_iX_j

are congruent to the OX_iY_j we may improve this to $n^2/4+cn$, but no further.

We now show that $g_6^{(2)}(n) \geqq c\, n^3$. Let $n \geqq 3$ be given. Let $n = 3m+r$, where $0 \leqq r < 3$. Choose a co-ordinate system in E_6, put $X_i = (a_i, b_i, 0,0,0,0)$ for $1 \leqq i \leqq m$, put $Y_i = (0,0,a_i,b_i,0,0)$ for $1 \leqq i \leqq m$, and put $Z_i =$ $(0,0,0,0,a_i,b_i)$ for $1 \leqq i \leqq m+r$, where (a_i,b_i) are $m+r$ distinct real solutions of $a^2 + b^2 = 1$. Then the $m^2(m+r)$ triangles $X_iY_jZ_k$ are all equilateral triangles of side length $\sqrt{2}$. Hence $g_6^{(2)}(n) \geqq m^2(m+r) \geqq c\, n^3$.

§5. Statement of the Main Theorems

Theorem 1 There exist n_1, $\varepsilon > 0$ such that $g_5^{(2)}(n) \leqq n^{3-\varepsilon}$ for $n \geqq n_1$. Consequently there exists a positive constant c such that $g_5^{(2)}(n) \leqq cn^{3-\varepsilon}$ for all n.

Let $|S|$ denote the cardinality of the set S. We shall deduce theorem 1 from the following theorem:

Theorem 2 Suppose A,B, and C are finite sets in E_5, such that $|A| \geqq M$, $|B| \geqq N$, and $|C| \geqq N$, where M and N are certain absolute constants. Then the triangles XYZ for X in A, Y in B, and Z in C cannot all have the same area, unless that area be zero.

§6. Some Graph Theory

By an r-graph $G^{(r)}$ we mean an object whose basic components are its elements, called vertices, and certain distinguished r-element sets of these elements, called r-sets. When r = 2, $G^{(r)}$ is an ordinary graph. When we say that G is a $G^{(r)}(n; m)$, we mean that G is an r-graph having n vertices and m r-sets. If G is a $G^{(r)}(n;\binom{n}{r})$, then G is the unique r-graph which has all possible r-element sets as its r-sets. We call this the complete r-graph on n vertices and denote it by $K^{(r)}(n)$. $K^{(r)}(n_1,\ldots,n_r)$ will denote

the r-graph of $n_1+\cdots+n_r$ vertices and $n_1\cdots n_r$ r-sets defined as follows: The vertices are

$$X_{i_j}^{(j)}, \quad 1 \leqq j \leqq r, \quad 1 \leqq i_j \leqq n_j$$

and the r-sets of our r-graph are the $n_1\cdots n_r$ r-sets

$$\{X_{i_1}^{(1)}, X_{i_2}^{(2)}, \ldots, X_{i_r}^{(r)}\}, \ 1 \leqq i_j \leqq n_j, \ 1 \leqq j \leqq r.$$

Denote by $f(n; K^{(r)}(\ell_1,\ldots,\ell_r))$ the smallest integer L so that every $G^{(r)}(n; L)$ contains a $K^{(r)}(\ell_1,\ldots,\ell_1)$.

In an elementary but non-trivial way Erdős [5] proves that if $n > n_o(r,\ell)$, then

(*) $\qquad f(n; K^{(r)}(\ell,\ldots,\ell)) \leqq n**(r - \ell**(1-r)).$

We shall use this result with r = 3, and we shall refer to the 3-sets of a 3-graph as triples in what follows.

§7. The Relation Between the Main Theorems

We now prove that theorem 2 implies theorem 1. Let ℓ be the maximum of M and N of theorem 2, let $\varepsilon=\ell^{-2}$, and let $X_1,\ldots,X_n$ be distinct points in E_5 with $n > n_o(r,\ell)$, where $n_o(r,\ell)$ is the function given in Erdős's inequality (*). It is an easy consequence of (*) that theorem 2 implies

(5) $$g_5^{(2)}(n; X_1,\ldots,X_n) \leqq n^{3-\varepsilon}.$$

To see this, let $\Delta > 0$, and let $G^{(3)}$ denote the 3-graph with n vertices $X_1,\ldots X_n$, where the triple $X_iX_jX_k$ is in $G^{(3)}$ if and only if the triangle $X_iX_jX_k$ has area Δ. Then theorem 2 implies that $G^{(3)}$ does not contain a $K^{(3)}(\ell,\ell,\ell)$ subgraph, and (5) then follows from (*). Theorem 1 follows, since Δ was arbitrary.

§8. Some Lemmas

Before proving theorem 2 we must introduce some definitions and lemmas. We shall use the notation $\{x\}$ to mean the least integer not less than x.

Lemma 1 Let triangles PX_iY_j, $1 \leq i \leq n+1$, $1 \leq j \leq H$ all have the same non-zero area Δ, where X_i, Y_j are points in real Euclidean n-dimensional space. If the n+1 distances $d(P,X_i)$ are all different and non-zero, then there are not more than 2^{n-1} distinct distances $d(P,Y_j)$. Hence at least $\{H/2^{n-1}\}$ of the Y_j are equidistant from P.

Proof Let P be the origin of co-ordinates. Let U_i be a unit vector parallel to $\vec{PX_i}$. The area of a triangle OXY can be written in terms of lengths and the inner product as half the square root of $|X|^2|Y|^2 - (X\cdot Y)^2$. For all i and j we have $4\Delta^2 = |X_i|^2|Y_j|^2 - (X_i\cdot Y_j)^2$, or $|Y_j|^2 - (U_i\cdot Y_j)^2 = r_i^2$, where $r_i = 2\Delta/|X_i|$. Let C_i be the set of solutions Y of

(6) $$|Y|^2 - (U_i\cdot Y)^2 = r_i^2.$$

In fact, C_i is a cylinder with axis U_i and radius r_i. Let k be the rank of the set $\{U_1,\ldots,U_{n+1}\}$. By renaming the U_i and choosing a suitable co-ordinate system, we may suppose that $U_i = (a_{i1},\ldots,a_{in})$ for $1 \leq i \leq n+1$, $a_{ii} \neq 0$ for $1 \leq i \leq k$, and $a_{ij} = 0$ if $j > k$ for all i. Putting $r = |Y|$ and $Y = (y_1,\ldots,y_n)$ in (6) and solving for $Y\cdot U_i$, we obtain

(7) $$\sum_{j=1}^{i} a_{ij}y_j = \pm\sqrt{r^2 - r_i^2} \qquad (1 \leq i \leq k)$$

$$\sum_{j=1}^{k} a_{k+1,j}y_j = \pm\sqrt{r^2 - r_{k+1}^2}.$$

We shall show that r^2 is the root of a non-zero polynomial of degree at most 2^{k-1}, and the lemma will follow. Let the system of equations $\sum_{j=1}^{i} a_{ij}y_j = z_i$ $(1 \leq i \leq k)$ have the solution $y_i = \sum_{j=1}^{i} b_{ij}z_j$ $(1 \leq i \leq k)$ and suppose

that $z_{k+1} = \sum_{j=1}^{k} a_{k+1,j}y_j$. Then substituting the expression for the y_j we get $z_{k+1} = \sum_{j=1}^{k} a_{k+1,j}\sum_{i=1}^{j} b_{ji}z_i = \sum_{j=1}^{k} c_j z_j$ for some c_j.

There are 2^k functions $f_i(t)$ of the form $\sqrt{t-r_{k+1}^2} - \sum_{j=1}^{k} \pm c_j\sqrt{t-r_j^2}$ corresponding to the choices of sign. Let $P(t) = f_1(t)\cdots f_m(t)$, where $m = 2^k$. If Y is a solution of (7), then clearly $P(r^2) = 0$. It is therefore sufficient to show that P is a non-zero polynomial of degree at most 2^{k-1}. Let $W_o = \sqrt{t-r_{k+1}^2}$ and let $W_i = c_i\sqrt{t-r_i^2}$ for $1 \leq i \leq k$. By induction on k we have $\Pi\ (W_o \pm W \pm\cdots\pm W_k) = F_k(W_o^2, W_1^2,\ldots,W_k^2)$, where F_k is a homogenous polynomial of degree 2^{k-1} and the product is taken over all 2^k possible combinations of signs. Hence $P(t) = F_k(t-r_{k+1}^2, c_1^2(t-r_1{}^2),\ldots,c_k^2(t-r_k^2))$ is a polynomial in t of degree at most 2^{k-1}.

To show that P is not the zero polynomial, we proceed as follows: $f_i(t) = \sqrt{t-r_{k+1}^2} - \sum_{j=1}^{k} \pm c_j\sqrt{t-r_j^2}$, $2f_i'(t) = 1/\sqrt{t-r_{k+1}^2} + \sum_{j=1}^{k} \pm c_j/\sqrt{t-r_j^2}$. Let $c_{k+1} = 1$, and let $R = r_p$ be the maximum r_j for which $c_j \neq 0$. Then $f_i'(t)$ is of constant sign for $R^2 < t < R^2 + \varepsilon_i$, some positive ε_i, since the term $c_p/\sqrt{t-r_p^2}$ goes to infinity as $t \downarrow R^2$, and the other terms remain bounded. (If the r_i were not distinct, considerable difficulty would arise at this point.) Hence f_i has at most one zero in that interval, and P has at most m zeros in the interval $R^2 < t < R^2 + \min \varepsilon_i$. The lemma is proved.

<u>Definitions</u> If all the points of a set B are equidistant from a point X, then we say that B is <u>equidistant from</u> X. If B is equidistant from every point X of A, then we say that B is <u>equidistant from</u> A. This relation is clearly not symmetric. If all the points of a set B are different distances from a point X, then we say that B is <u>separated by</u> X. If B is

separated by every point X of A, then we say that B is <u>separated by</u> A. This relation is also not symmetric.

<u>Lemma 2</u> Let S and T be arbitrary sets of cardinalities M and N respectively and suppose that the elements of S × T are divided into two classes C_1 and C_2. (Suppose that the pairs are colored two colors C_1 and C_2.) Then there is a subset T' of T of cardinality $\{N/(2^M)\}$ such that for every X in S, the elements (X,Y) for Y in T' are either all in C_1 or all in C_2. (For every X, the color of the pair (X,Y) for Y in T' depends only on X.)

<u>Proof</u> Use induction on M and the pigeon-hole principle.

<u>Lemma 3</u> Given pairwise disjoint finite subsets A, B, C of E_k, there are subsets A' of A, B' of B, and C' of C such that B' is separated by or equidistant from A' and C' is separated by or equidistant from A'. Further if $|A| = H$, we have $|A'| = \{H/4\}$, $|B'| = |B|{**}2^{-H}$, and $|C'| = |C|{**}2^{-\{H/2\}}$.

<u>Proof</u> Let $B_o = B$, and $i \geq 1$. If the elements of A are $X_1, X_2, \ldots, X_H$, we define sets $B_1, B_2, \ldots B_H$ as follows. For each X_i we color X_i and take a subset of B_{i-1} as follows. If B_{i-1} has a subset of $\{\sqrt{|B_{i-1}|}\}$ points separated by X_i, let B_i be this subset, and we color X_i red. Otherwise, by the pigeon-hole principle there is a subset B_i of B_{i-1} of cardinality $\{\sqrt{|B_{i-1}|}\}$, equidistant from X_i, and we color X_i blue. If we do this for $1 \leq i \leq H$, we get a subset B_H of B of cardinality $|B|{**}2^{-H}$ that is separated by all the red X_i and equidistant from all the blue X_i. Let $B' = B_H$. Then there is a subset A* of cardinality {H/2} of A such that B' is either separated by or equidistant from A^*.

Similarly there is a subset C' of C of cardinality $|C|{**}2^{-K}$, $K = \{H/2\}$, and a subset A' of A* cardinality {H/4} such that C' is either separated by or equidistant from A'. The lemma follows.

Lemma 4 Let PX_iY_j have area $\Delta > 0$ for $1 \leqq i \leqq 3$ and $1 \leqq j \leqq 3$, where X_1, X_2, X_3 are distinct and Y_1, Y_2, Y_3 are distinct. Then the X_i are not collinear. By symmetry, the Y_j are not collinear.

Proof Suppose the X_i are collinear. Let $1 \leqq j \leqq 3$. The X_i lie on the surface of a cylinder with axis PY_j. This can happen only if the X_i lie on a line parallel to PY_j. Consequently, P and the Y_j are collinear. The distances $d(P,Y_j)$ cannot all be equal. Suppose, without loss of generality, that $d(P,Y_1) > d(P,Y_2)$. Then triangle PX_1Y_1 has a greater area than triangle PX_1Y_2, contrary to the hypothesis.

Lemma 5 If in E_4 the cylinder $C = \{X: |X|^2 - (U\cdot X)^2 = c^2\}$, where $|U| = 1$, intersects the hyperplane π, then there are three possibilities.

(i) If $\vec{OU}$ is parallel to π, then C intersects π in a cylinder.

(ii) If $\vec{OU}$ is perpendicular to π, then C intersects π in a sphere.

(iii) If neither of the above, then C intersects π in an ellipsoid of revolution whose axis is the projection of $\overleftrightarrow{OU}$ onto π.

Proof Choose the origin O to be on π and choose the X_4 axis normal to π. Then π is the set of points (x_1,x_2,x_3,x_4) such that $x_4=0$. Now choose the X_1 axis lying in π, in the direction of the projection $\vec{OU}^*$ of $\vec{OU}$ onto π. Then $U = (\alpha,0,0,\beta)$ where $\alpha^2 + \beta^2 = 1$. The cylinder C has the equation $\sum_1^4 x_i^2 - (\alpha x_1 + \beta x_4)^2 = c^2$ and C intersects π in a surface with equation

$\beta^2 x_1^2 + x_2^2 + x_3^2 = c^2$. If $\beta = 0$ we have (i), if $\beta = 1$, we have (ii), and if $0 < \beta < 1$, then we have (iii).

§9. Proof of Theorem 2

Let A,B,C be sets of cardinality M,N,N respectively such that the triangles XYZ for X in A, Y in B and Z in C all have a common positive area Δ. We shall show that this leads to a contradiction if M and N are large

enough, and the theorem will follow.

Let us assume that $N \geqq M \geqq 3$; then, by lemma 4 no three points of A (or B or C) are collinear. By lemma 3 and the symmetry between B and C, we may suppose without loss of generality that one of the following holds:

(I) B is separated by A, but C is equidistant from A.

(II) B is separated by A and C is separated by A.

(III) B is equidistant from A, and C is equidistant from A.

This application of lemma 3 reduces M and N. From now on, M and N are for the new sets.

Let $A = \{X_1, \ldots, X_M\}$, $B = \{Y_1, \ldots, Y_N\}$. First of all, (II) leads immediately to a contradiction. Take one point X of A and six points $Y_1, \ldots, Y_6$ of B. Then by lemma 1, there are at most 16 points $Z_1, \ldots, Z_{16}$ such that $\{Z_1, \ldots, Z_{16}\}$ is separated by X. We get a contradiction if $N \geqq 17$.

Secondly, (III) leads to a contradiction; (III) implies that the affine hull of A is orthogonal to the affine hull of B and the affine hull of C. We shall find subsets B' of B and C' of C whose affine hulls are orthogonal. Let X be a fixed point of A, let B' be a subset of order three of B, let d be the common distance of B from X, and let e be the common distance of C from X. Then $4\Delta^2 = e^2d^2 - \{(Y-X)\cdot(Z-X)\}^2$ or $|(Y-X)\cdot(Z-X)| = \sqrt{e^2d^2-4\Delta^2}$ for all (Y,Z) in $B' \times C$. By lemma 2 there is a subset C' of C of order $\{N/8\}$ such that for each Y in B', $(Y-X)\cdot(Z-X)$ has a constant sign as Z ranges over C'. Hence for Z, Z' in C' and Y, Y' in B', $(Y-X)\cdot(Z-X) = (Y-X)\cdot(Z'-X)$, $(Y-X)\cdot(Z-Z') = 0$, $(Y-Y')\cdot(Z-Z') = 0$, and the affine covers of C' and B' are orthogonal. Since no three points of A (or B or C) can be collinear, and three pairwise orthogonal planes cannot exist in R^5, we obtain a contradiction for $M \geqq 3$ and $N \geqq 17$.

We next show that (I) leads to a contradiction. This is the last and

hardest case. We start by reducing B to be $\{Y_1,\ldots,Y_M\}$ by throwing away N-M points. Now

$4\Delta^2 = |X-Z|^2|Y-X|^2-\{(Y-X)\cdot(Z-X)\}^2$ for all (X,Y,Z) in $A \times B \times C$. Hence $|(Z-X)\cdot(Y-X)| = \sqrt{|X-Z|^2|Y-X|^2-4\Delta^2}$, and the right hand side is independent of Z, since $|X-Z|$ is independent of Z. Let $\gamma_1,\cdots,\gamma_r$ where $r = M^2$, be an enumeration of $A \times B$. Let us 2-color $G^{(2)} = (A \times B) \times C$ as follows: If $(Z-X)\cdot(Y-X) \geqq 0$ then color ((X,Y),Z) red. Otherwise color ((X,Y),Z) blue. By lemma 2, there is a subset C' of C of cardinality $\{N/(2^r)\}$ such that $(Z-X)\cdot(Y-X)$ is of constant sign as Z ranges over C' with (X,Y) fixed. Hence for (X,Y) in $A \times B$ and Z, Z' in C', $(Z-Z')\cdot(Y-X) = 0$; for Y,Y' in B, Z,Z' in C' and X,X' in A we have $(Z-Z')\cdot(Y-Y') = (Z-Z')\cdot(X-X') = 0$. Hence the affine hull of C' is orthogonal to the affine hull of $A \cup B$. Let us assume that $N \geqq 2^{**}(M^2+1)+1$, so that the order, $\{N/(2^r)\}$, of C' is at least 3, so that C' contains three non-collinear points.

Hence the dimension of the affine hull of C' is at least two, and this forces $A \cup B$ to lie in a three dimensional subspace π.

If C' is also contained in π, then the whole configuration is in R^3 and if M and N are large enough, we have a contradiction by (4). We may therefore suppose the existence of a point Z of C' that is not in π. Let Z* be the orthogonal projection of Z onto π. The points of B lie on cylinders with axes $\overleftrightarrow{X_iZ}$ $(1 \leqq i \leqq M)$, which by lemma 5 intersect π in surfaces ξ_i which are either cylinders, spheres, or ellipsoids of revolution with axes $\overleftrightarrow{X_iZ^*}$. Call these surfaces ξ_i. By the same lemma, ξ_i cannot be a cylinder, since $\overleftrightarrow{X_iZ}$ is not parallel to π.

Also by lemma 5, ξ_i is a sphere only if $X_i=Z^*$. Since no three points of the set A are collinear, there exist two points, say X_1 and X_2, of A such that X_1,X_2,Z^* are not collinear. This implies that neither X_1 nor X_2

coincide with Z^*, so that by lemma 5, ξ_1 and ξ_2 are ellipsoids of revolution with axes of revolution $\overleftrightarrow{X_1Z^*}$ and $\overleftrightarrow{X_2Z^*}$.

Suppose that B has a nine-point subset B^* that is equidistant from Z. Then B^* lies on a sphere S^* having center Z^* and lying in π, and for i=1,2 each ξ_i intersects the sphere S^* in a pair of circles C_i and C_i' whose centers lie on the line $\overleftrightarrow{X_iZ^*}$. For i,j=1 or 2 and $j \neq i$, C_i is distinct from C_j and C_j', since the normals $\overleftrightarrow{X_iZ^*}$ and $\overleftrightarrow{X_jZ^*}$ are not parallel, due to the fact that X_1, X_2, Z^* are non collinear. Two distinct circles on the surface of a sphere in R^3 intersect in at most two points. Hence $(C_1 \cup C_1') \cap (C_2 \cup C_2') = C_1 \cap C_2 \cup C_1' \cap C_2 \cup C_1 \cap C_2' \cup C_1' \cap C_2'$ is a set of order less than nine containing a set B^* of order nine, which is absurd.

Hence there exists a set B' of cardinality $\{M/8\}$, which is separated by Z. Let us suppose that $M \geq 41$ and take B' to have cardinality at least 6. By lemma 1 there exists a subset A' of A of cardinality $R = \{M/16\}$ that is equidistant from Z; let d be the common distance of the points of A' from Z. Let B" be a subset of three elements of B'. Then $4\Delta^2 = d^2|Y-Z|^2 - \{(Y-Z)\cdot(X-Z)\}^2$ or $|(Y-Z)\cdot(X-Z)| = \sqrt{d^2|Y-Z|^2-4\Delta^2}$ for all $(X,Y) \in A' \times B''$. The right hand side is independent of X. By lemma 2, there is a subset A" of A of order $S = \{R/8\}$ such that $(Y-Z)\cdot(X-Z)$ is of constant sign for fixed Y in B" as X ranges over A". Hence $(Y-Z)\cdot(X-X') = 0$ for all X, X' in A" and Y in B". Hence $(Y-Y')\cdot(X-X') = 0$ for all X, X' in A" and Y, Y' in B", and the affine hull of A" is orthogonal to the affine hull of B". Combining this with our earlier result, we see that the affine hulls of A", B", and C' are pairwise orthogonal. To get a contradiction, it is sufficient to ensure that each of these sets has at least three elements. If M = 257, then R = 17 and S = 3. If $N \geq 2^{**}(M^2+1)+1$ we obtain the desired contradiction.

REFERENCES

[1] P. Erdős, On sets of distances of n points, Amer. Math. Monthly, 53(1946), 248-250.

[2] L. Moser, On the different distances determined by n points, Amer. Math Monthly, 59(1952), 85-91.

[3] P. Erdős, On some applications of graph theory to geometry, Canadian Journal Math., 19(1967), 968-971, see also P. Erdős, On sets of distances of n points in euclidean space, Publ. Math. Inst. Hung. Acad. Sci., 5(1960), 165-169.

[4] T. Kovari, V. T. Sos and P. Turan, On a problem of K. Zarankiewicz, Coll. Math., 3(1945), 50-57.

[5] P. Erdős, On extremal problems of graphs and generalized graphs, Israel J. Math., 2(1964), 183-190, Theorem 1.

[6] P. Erdős and G. Purdy, Some extremal problems in geometry, Journal of Combinatorial Theory, 10(1971), 246-252.